

1982

Studies on Several Biologically Active Peptides and Synthetic Random Copolypeptides.

Randall Murdock Robinson

Louisiana State University and Agricultural & Mechanical College

Follow this and additional works at: https://digitalcommons.lsu.edu/gradschool_disstheses

Recommended Citation

Robinson, Randall Murdock, "Studies on Several Biologically Active Peptides and Synthetic Random Copolypeptides." (1982). *LSU Historical Dissertations and Theses*. 3739.
https://digitalcommons.lsu.edu/gradschool_disstheses/3739

This Dissertation is brought to you for free and open access by the Graduate School at LSU Digital Commons. It has been accepted for inclusion in LSU Historical Dissertations and Theses by an authorized administrator of LSU Digital Commons. For more information, please contact gradetd@lsu.edu.

INFORMATION TO USERS

This was produced from a copy of a document sent to us for microfilming. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help you understand markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure you of complete continuity.
2. When an image on the film is obliterated with a round black mark it is an indication that the film inspector noticed either blurred copy because of movement during exposure, or duplicate copy. Unless we meant to delete copyrighted materials that should not have been filmed, you will find a good image of the page in the adjacent frame. If copyrighted materials were deleted you will find a target note listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed the photographer has followed a definite method in "sectioning" the material. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For any illustrations that cannot be reproduced satisfactorily by xerography, photographic prints can be purchased at additional cost and tipped into your xerographic copy. Requests can be made to our Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases we have filmed the best available copy.

University
Microfilms
International

300 N. ZEEB RD., ANN ARBOR, MI 48106

Robinson, Randall Murdock

**STUDIES ON SEVERAL BIOLOGICALLY ACTIVE PEPTIDES AND
SYNTHETIC RANDOM COPOLYPEPTIDES**

The Louisiana State University and Agricultural and Mechanical Col. **PH.D. 1982**

**University
Microfilms
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy.
Problems encountered with this document have been identified here with a check mark ✓.

1. Glossy photographs or pages _____
2. Colored illustrations, paper or print _____
3. Photographs with dark background _____
4. Illustrations are poor copy _____
5. Pages with black marks, not original copy _____
6. Print shows through as there is text on both sides of page _____
7. Indistinct, broken or small print on several pages ✓
8. Print exceeds margin requirements _____
9. Tightly bound copy with print lost in spine _____
10. Computer printout pages with indistinct print _____
11. Page(s) _____ lacking when material received, and not available from school or author.
12. Page(s) _____ seem to be missing in numbering only as text follows.
13. Two pages numbered _____. Text follows.
14. Curling and wrinkled pages _____
15. Other _____

University
Microfilms
International

STUDIES ON SEVERAL BIOLOGICALLY ACTIVE PEPTIDES AND
SYNTHETIC RANDOM COPOLYPEPTIDES

A Dissertation

Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
In partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

in

The Department of Biochemistry

by

Randall M. Robinson
B.S., Middle Tennessee State University, 1977
May 1982

ACKNOWLEDGEMENT

The author wishes to express deep gratitude to Dr. Wayne L. Mattice for patient guidance and support. Gratitude is also extended to Drs. E.S. Younathan, R.S. Allen, and D.K. Carpenter for serving on the advisory committee. I also wish to thank Dr. E.W. Blakeney for both serving on my advisory committee and also for crucial experimental support. I wish to thank Dr. J. Skolnick for stimulating discussions of portions of this work. I also wish to thank Dr. A. Holltzer in absentia for help in the form of related dissertations provided to me by Dr. Skolnick.

Aid in various forms from Dr. Erin Hawkins and German Santiago is also greatly appreciated. I also wish to thank my wife, Charlene, for her support and help with the final typing.

Financial support from the Biochemistry Department and the National Science Foundation is genuinely appreciated.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	Page ii
LIST OF TABLES	Page iv
LIST OF FIGURES	Page v
ABSTRACT	Page viii
CHAPTER 1 - INTRODUCTION	Page 1
PART ONE - CONFORMATIONAL STUDIES OF BIOLOGICALLY ACTIVE PEPTIDES	Page 44
CHAPTER 2 - CONFORMATIONAL PROPERTIES OF HUMAN BETA-ENDORPHIN AND BETA-LIPOTROPIN IN WATER AND IN THE PRESENCE OF ANIONIC LIPIDS	Page 45
CHAPTER 3 - LIPID INDUCED CONFORMATIONAL CHANGES IN GLUCAGON, SECRETIN, AND VASOACTIVE INTESTINAL PEPTIDE	Page 72
CHAPTER 4 - β -SHEET TO α -HELIX TRANSITION IN THE BINDING SUBUNIT OF CHOLERA TOXIN	Page 93
PART TWO - STUDIES ON SYNTHETIC RANDOM COPOLYPEPTIDES	Page 114
CHAPTER 5 - HELIX INITIATION AND PROPAGATION BY HYDROXYETHYL- L-GLUTAMINYL RESIDUES IN WATER	Page 115
CHAPTER 6 - ELECTROSTATIC INTERACTIONS IN RANDOM COPOLY- PEPTIDES CONTAINING L-GLUTAMYL RESIDUES	Page 134
BIBLIOGRAPHY	Page 160
VITA	Page 171

LIST OF TABLES

Table 1 - Statistical Weights for Possible Conformations of a Tripeptide	Page 33
Table 2 - Statistical Weights Used For Amino Acid Residues in Water	Page 40
Table 3 - Parameters for Random Coil to Helix Transition in Water at 20°C for the ω -Hydroxyalkyl-L-glutaminyl Residues	Page 129

LIST OF FIGURES

- Figure 1 - A Detailed Representation of a Portion of a
Polypeptide Chain Page 3
- Figure 2 - The Ramachandran Steric Map for the L-Alanyl
Residue Page 7
- Figure 3 - The Conformational Energy Map For the L-Alanyl
Residue Page 10
- Figure 4 - Cartesian Reference Frames For Bonds of Chain
Molecules Page 13
- Figure 5 - The Alpha-Helix and the Disposition of the
Sidechains in the Alpha-Helix Page 28
- Figure 6 - The Amino Acid Sequences for Human, Ovine, and
Porcine β -Lipotropin Page 48
- Figure 7 - Helix Propagation Probability Profiles for
Human β -Endorphin Page 63
- Figure 8 - Proposed Helical Segment in Human β -Endorphin ... Page 65
- Figure 9 - A Plot of Binding Activity Versus Computed Helical
Content in SDS for β -Endorphin Fragments Page 70
- Figure 10 - The Amino Acid Sequences of Glucagon, Secretin,
and Vasoactive Intestinal Peptide(VIP) Page 74

Figure 11 - Circular Dichroism(CD) Spectra for Glucagon, Secretin, and VIP in Water and SDS	Page 80
Figure 12 - CD Spectra of Glucagon, Secretin, and VIP in Phosphatidic Acid	Page 84
Figure 13 - Helix Propagation Probability Profiles for Glucagon, Secretin, and VIP	Page 88
Figure 14 - A Depiction of the Quaternary Structure for the Binding Subunit of Cholera Toxin	Page 95
Figure 15 - The Amino Acid Sequence for the Binding Subunit of Cholera Toxin	Page 98
Figure 16 - Circular Dichroism of the Binding Subunit in Water and in SDS	Page 102
Figure 17 - Time Dependence of $[0]_{222}$ for Subunit B in SDS ..	Page 104
Figure 18 - Helix Propagation Probability Profile for the Binding Subunit of Cholera Toxin	Page 108
Figure 19 - The Proposed Helical Hairpin Structure of the B Subunit in SDS	Page 111
Figure 20 - Experimentally Determined Fraction Helix as a Function of Hydroxybutyl-L-glutamine Composition	Page 123
Figure 21 - The Parameter s as a Function of Temperature	Page 125
Figure 22 - A Plot of $\ln s$ Versus $1/T$	Page 130

Figure 23 - A Typical Titration Curve of a Nonrigid Polyelectrolyte	Page 138
Figure 24 - Titration Curves for the L-Glutamic Acid- L-Tyrosine Copolymer	Page 147
Figure 25 - Titration Curves for the L-Glutamic Acid- L-Alanine Copolymer	Page 149
Figure 26 - A Plot of Fraction Helix and Alpha as a Function of pH for the L-Glutamic Acid-L-Tyrosine Copolymer	Page 153
Figure 27 - A Plot of Fraction Helix and Alpha as a Function of pH for the L-Glutamic Acid-L-Alanine Copolymer	Page 155

ABSTRACT

The work presented in this dissertation is based on the Zimm-Bragg model for the helix-coil transition in polypeptides. The molecules studied are of two distinct types ; biologically active peptides and synthetic random copolypeptides. Biologically active peptides discussed include : β -endorphin, β -lipotropin, glucagon, secretin, vasoactive intestinal peptide, and the binding or B subunit of cholera toxin. Random copolypeptides studied included copolymers of ω -hydroxyethyl and ω -hydroxybutyl-L-glutaminy residues, and ionic copolypeptides containing L-glutamyl and either L-alanyl or L-tyrosyl residues.

The studies on biologically active polypeptides involved rationalizing the interactions of the aforementioned peptides with anionic lipids. The model used is based on the assumption that interaction of these peptides with anionic lipids increased the tendency for helix formation by histidyl, arginyl, and lysyl residues. The model allows for the even distribution of the end effects to both ends of a helical segment. Another important aspect of the calculation is that the configuration partition function can be used to predict the portion of the chain with the highest probability for helix formation in the presence of anionic lipids.

Residues 17-29 of β -endorphin are expected from the calculation to be alpha helical in the presence of anionic lipids. This portion of the chain forms an amphipathic helix which would be expected to interact favorably with a membrane containing anionic lipids.

The order of helix formation is predicted to be glucagon < secretin < vasoactive intestinal peptide, which is in agreement with experiment-

al circular dichroism measurements. The region of highest helix propagation probability is residues 13-20. This prediction is in harmony with various experiments by other workers, and this segment of the chain has been implicated in receptor binding for secretin and vasoactive intestinal peptide.

A helical hairpin structure is predicted for the B subunit of cholera toxin in dodecyl sulfate. The transition from β -structure to α -helix occurs over a length of time comparable to the previously described lag period observed in vivo and in vitro. The conformational transition may be related to the pathogenesis of the toxin.

The host guest technique was utilized in order to determine σ and s for ω -hydroxyethyl-L-glutamyl residues in water. Hydroxybutyl-L-glutamyl served as the host residue. A temperature independent value of 1×10^{-5} for σ . The value of s was seen to vary from 0.940 at 4°C to 0.925 at 64°C. The results here are consistent with a smooth decrease in such values from hydroxybutyl to hydroxypropyl to hydroxyethyl-L-glutamyl residues. These results would rule out the possibility of several theoretical structure for this residue e.g. intraresidue hydrogen bonding, at least in a polar solvent such as water.

The Zimm-Rice theory for the helix-coil transition in charged homopolypeptides was used in an attempt to rationalize circular dichroism and potentiometric titration data for random copolymers containing L-glutamyl residues along with either L-alanyl or L-tyrosyl residues. This method was found to fail and no rationalization could be given for adjusting the parameters of the model to bring about agreement. One possible reason may be the model's failure to treat 1-3 and 1-4

interactions.

CHAPTER 1 INTRODUCTION

The body of this work is involved in determining to what extent the local structure of relatively large chain molecules determines the dimensions and other properties of such molecules. The physical chemical principles are the same whether the molecule in question is polymethylene, a polymer of commercial importance, or β -endorphin, a specific sequence polypeptide of biochemical, pharmacological, and medical interest. Each type of molecule has distinctive distinguishing characteristics, and the task which is attempted in this introduction is to determine what assumptions are valid, and also which methods represent a cogent analysis of the characteristics of a given molecule. This chapter is a self-contained introduction to the basic statistical mechanical machinery used to rationalize certain conformational properties of polypeptides. The chapters which follow are studies which range in scope from very specific biologically important interactions between proteins and lipids to basic studies designed to yield fundamental information pertinent to many varied situations.

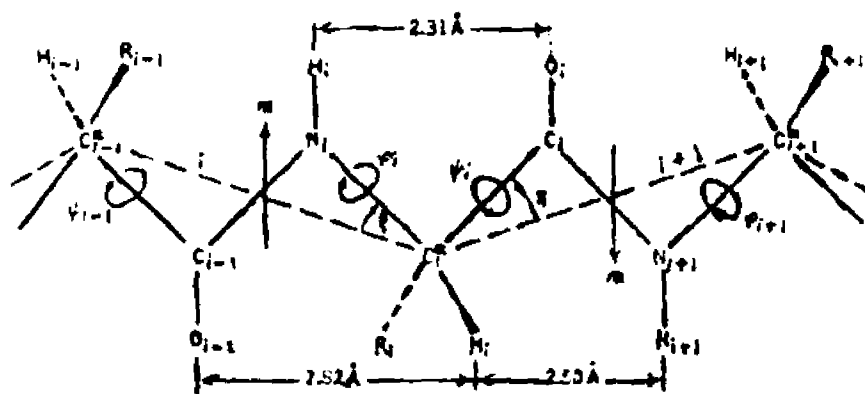
One very formidable problem which has not been solved is how three dimensional information is stored and transmitted via the genetic code. This code determines the sequence of the amino acid residues which make up a given protein molecule. Modern biophysical research has shown that the sequence of the amino acid residues determines the "native" structure or conformation which the protein adopts.¹ The question of how this sequence determines which conformation among the staggering number of conformations accessible to the chain it will adopt as the native conformation has not yet been completely answered. However,

great strides have been made towards this end.

The structural properties of proteins have been divided into several classes or hierarchies. These classes are primary, secondary, tertiary, quaternary, and pentiary structure.² The primary structure of a protein is the sequence of the amino acid residues which make up the molecule commencing with the amino terminal residue (N-terminus), and ending with the carboxyl terminal residue (C-terminus). Secondary structure is the name given to any regular arrangement of adjacent residues in space. Tertiary structure refers to the folding of these secondary structures into the native globular or fibrous conformation. Quaternary structure involves the interactions between component chains of oligomeric proteins. Pentiary structure is the most complex of all, and refers to the formation of large complexes by several large oligomeric proteins. A good example of this is the pyruvate dehydrogenase complex, which is formed by several large enzyme molecules.³

Before considering any specific conformations of a peptide, it is important to first consider the structure of the peptide backbone and the implications for the conformational properties of proteins. A representative portion of a polypeptide backbone is shown in figure 1. Due to the partial double bond character of the amide bonds, these bonds remain for the most part in the planar trans configuration,⁴ with the exception of certain proline residues.⁵⁻⁸ This means that six atoms, namely, the amide nitrogen and hydrogen, the carbonyl carbon and oxygen, and two successive alpha carbons are confined to a plane. As a consequence of this fact, neighboring alpha carbons are equidistant to a good approximation. This means that a polypeptide chain can be considered as a chain of alpha carbons connected by virtual bonds as

Figure 1. A detailed representaion of a portion of a polypeptide chain in the fully extended conformation for which, by convention,⁹ the values of ϕ and ψ are equal to zero. Successive amino acid residues are indexed $i-1$, i , and $i+1$. Virtual bonds are shown as dashed lines. Group dipole moments are indicated by arrows (μ). $\eta = 22.2$, $\xi = 3.2$, and l_i , the length of a virtual bond, is equal to 3.8 Angstroms. ¹⁰



shown in figure 1. With these restrictions on the conformational properties of the polypeptide backbone it is of interest to note exactly which variables determine the conformation of the backbone. The variables which fit this description are the two rotational angles ϕ and ψ . When the values for these two angles are known the conformation of the backbone can be rigorously described. Unfortunately, there are two conventions used in the literature to describe the reference state of ϕ and ψ .^{9,11} The convention used in this dissertation takes the fully extended chain as the reference state, and thereby assigns a value of zero for the angles ϕ and ψ for this particular conformation.⁹

The values for these angles for succeeding residues may be identical as is the case with certain secondary structures, or they may be distributed randomly within certain limits as is the case for a random coil polypeptide. The term random coil here refers to a concept from polymer chemistry.¹⁰ This specifically means a chain molecule which does not adopt any specific conformation, but has conformational properties which are random, and the properties of the chain can be rationalized best via statistical mechanics. Many synthetic homopolypeptides and biologically active peptides exist as random coils in aqueous solution. Therefore it is of great importance to study and try to rationalize the properties of the random coil polypeptide. When insight is gained into the random coil structure for a given chain, information about which conformations are accessible to the chain, and also which ones are more favorable from an energetic standpoint, and may therefore be prevalent is also gained.

In order to assess which conformations are possible, and to compare them from an energetic standpoint, the potential energy as a function

of ϕ and ψ must be determined. The first step in this direction was taken by Ramachandran and his coworkers,¹² who used a hard sphere potential in order to determine which conformations could be ruled out on steric grounds. They used this potential to distinguish allowed from disallowed regions of conformational space. Then they reduced the interaction radii, and designated another region as partially allowed. Figure 2 shows the Ramachandran map for an L-alanyl residue. The location of several ordered structures are noted on this figure. It is of interest at this point to note that all of these structures appear in either allowed or partially allowed regions of the map.

This same idea was explored in a much more rigorous manner by P. J. Flory who used more realistic potential functions to describe the local interactions within the peptide chain. The expression used for calculating the conformational energy is as follows ; ^{13,14,15}

$$E(\phi, \psi) = (E_{\phi}^0 / 2) (1 - \cos 3\phi) + (E_{\psi}^0 / 2) (1 - \cos 3\psi) + \sum_{k,l} E_{kl}(\phi, \psi) + E_c \quad (1)$$

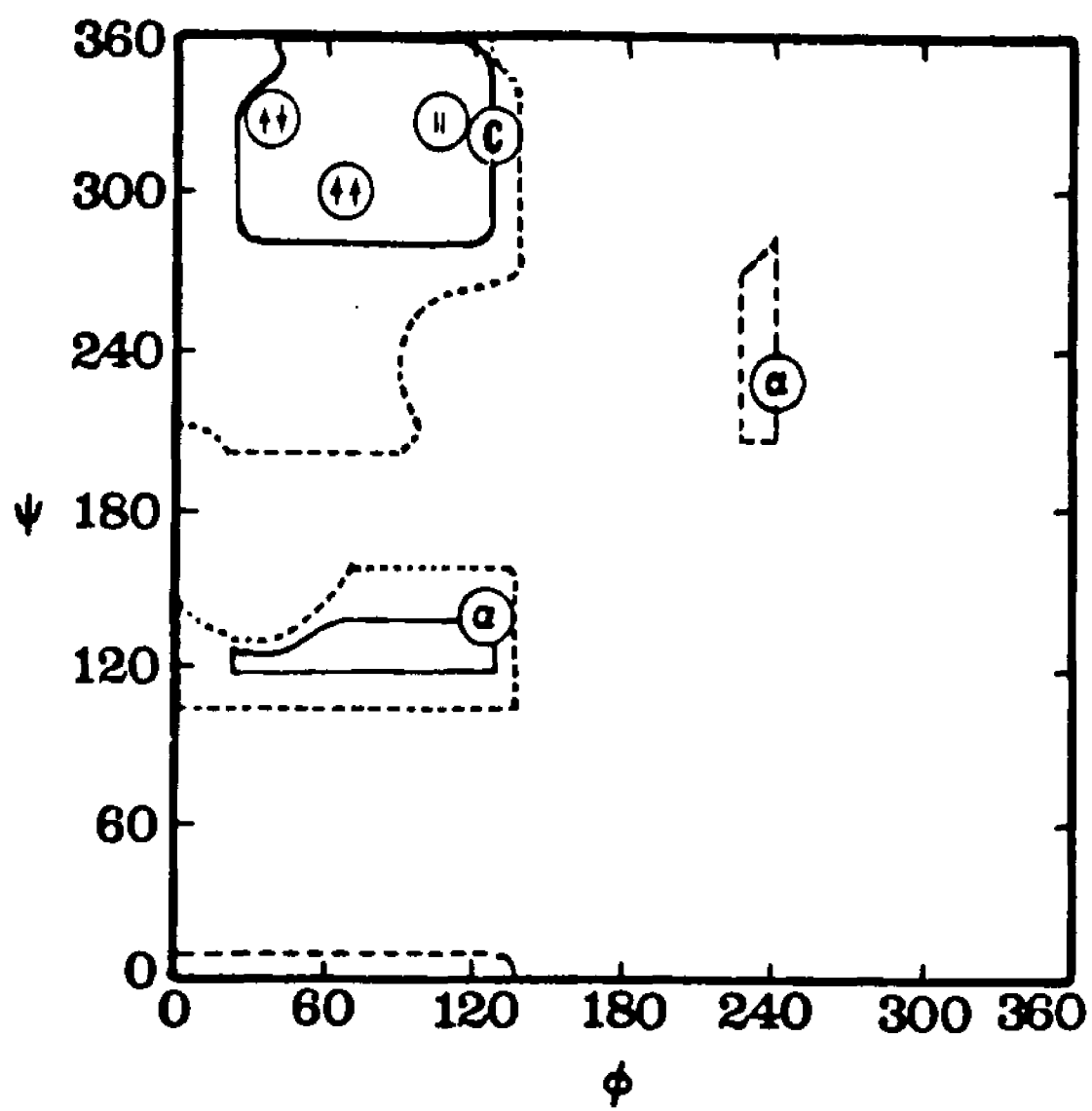
The summation term includes the interactions for every nonbonded pair of atoms whose distance of separation is dependent upon the value of ϕ and ψ . The E_c term represents the coulombic interactions which arise from the electrical asymmetries of the chain caused by the large dipole moments of the adjacent amide groups. The interaction between atoms k and l can be calculated by the following expression ;

$$E_{kl}(\phi, \psi) = a_{kl}/r_{kl}^m - c_{kl}/r_{kl}^6 \quad (2)$$

where the first term represents repulsive and the second term

Figure 2. The Ramachandran steric map ¹⁶ for the L-alanyl residue.

Allowed regions are enclosed by solid lines; partially allowed regions are enclosed by dashed lines. Conformations for several common ordered structures are denoted on the map as follows; right and left handed alpha helices, α , parallel pleated sheet, $\uparrow\uparrow$, antiparallel pleated sheet, $\uparrow\downarrow$, polyglycine -II and poly-L-proline-II, II ,and the collagen helix, C .



attractive interactions. When the exponent m has a value of 12, then the potential is the familiar Lennard-Jones potential.¹⁷ When a calculational method such as this is used there are no cutoff points or allowed or disallowed regions, instead a continuous potential energy function is obtained. The conformational energy map for the L-alanyl residue calculated in this manner is depicted in figure 3.

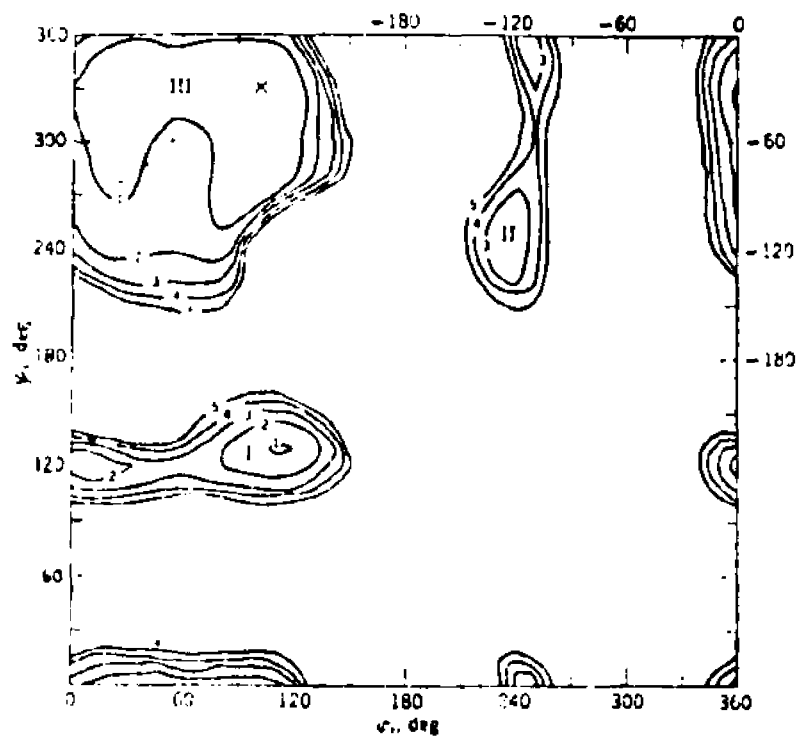
These calculations have shed a great deal of light upon those interactions which are important to the conformational energies of polypeptides. More importantly, however, this information can be used in conjunction with other methods to calculate quantities which are experimentally accessible. This is, of course, where the true test of any theoretical calculation lies. Among the quantities whose calculation is outlined in this chapter are ; the end to end vector, \underline{r} , the square of the magnitude of \underline{r} , r^2 , the radius of gyration, s , and the square of the radius of gyration. The calculation of the statistical mechanical averages of these quantities and also the calculation of the characteristic ratio is also outlined. Other quantities such as the dipole moment and optical anisotropy can be calculated in a similar manner.¹⁸

The end to end vector for a chain molecule in a particular conformation may be written as follows ;

$$\underline{r} = \sum_{i=1}^n \underline{l}_i \quad (3)$$

where \underline{l}_i is the virtual bond expressed as a vector, and all of the virtual bond vectors are expressed in the same reference frame. The way in which the coordinate system for a given bond is designated

Figure 3. The conformational energy map for the L-alanyl residue calculated using equation 1, with E_c estimated on the basis of point monopole charges. The contour lines represent a change of 1 Kcal mole⁻¹. The energy minimum is denoted by an x. 15



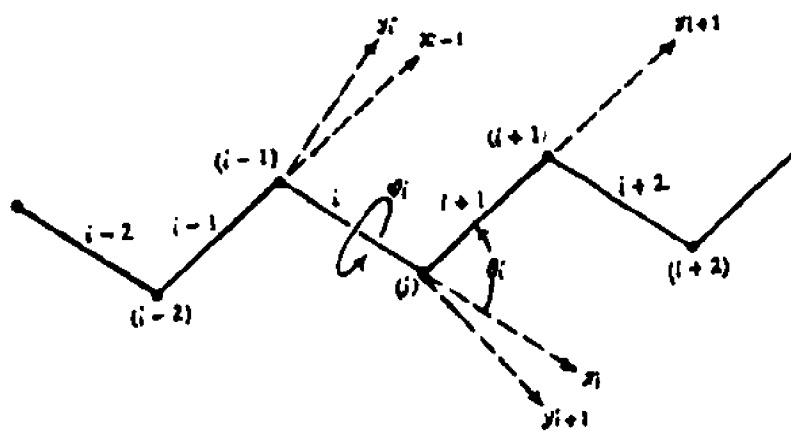
is shown in figure 4. For a polypeptide chain this would represent the alpha carbons connected by virtual bonds. It is at once obvious from the discussion above that in order to calculate the end to end vector, a method must be found for expressing a given bond vector in one reference frame in the reference frame of another bond vector.

The mathematical method employed for expressing a given vector in the coordinate system of the preceeding bond vector is the similarity transformation. It can be shown that exists a matrix, T_i , which is capable of transforming a vector y into the coordinate system of vector y' . For a simple bond as shown in figure 4 this matrix can be written ;

$$T_i = \begin{bmatrix} \cos\theta_i & \sin\theta_i & 0 \\ \sin\theta_i \cos\phi_i & -\cos\theta_i \cos\phi_i & \sin\phi_i \\ \sin\theta_i \sin\phi_i & -\cos\theta_i \sin\phi_i & -\cos\phi_i \end{bmatrix} \quad (4)$$

where $y' = T_i y$. The actual transformation of a virtual bond in a polypeptide chain from one reference frame to another is not quite as simple as the process outlined above. This process actually involves three separate transformations, and therefore three transformation matrices. In the process of transformation, the former reference frame is brought into coincidence with the final one by a rotation τ about the z axis of the initial reference frame, followed by a rotation ρ about the x axis of the final reference frame. The axes $(x y z)_{i+1}$ are related to the reference frame of the actual bond $C^\alpha-C$ of residue i by the following rotations ; $\tau = -\eta$, and $\rho = -\psi_i$. Therefore the transformation matrix for this process may be written as ;

Figure 4. Cartesian reference frames as defined for the bonds of a chain molecule. The z axes are not shown. They are perpendicular to the plane of the figure, and are designated so as to complete a right-handed coordinate system.¹⁰



$$T(-\eta, -\psi) = \begin{bmatrix} \cos\eta & \sin\eta & 0 \\ -\sin\eta\cos\psi_i & \cos\eta\cos\psi_i & -\sin\psi_i \\ -\sin\eta\sin\psi_i & \cos\eta\sin\psi_i & \cos\psi_i \end{bmatrix} \quad (5)$$

The transformation from bond $C^\alpha-C$ to bond $C^\alpha-N$ makes use of the rotations $\tau = \theta$, and $\rho = \pi - \phi_i$, and the subsequent transformation matrix is ;

$$T(\theta, \pi - \phi_i) = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ \sin\theta\cos\phi_i & -\cos\theta\cos\phi_i & \sin\phi_i \\ \sin\theta\sin\phi_i & -\cos\theta\sin\phi_i & -\cos\phi_i \end{bmatrix} \quad (6)$$

The transformation is completed by transforming from $C^\alpha-N$ to virtual bond i . This involves one rotation, namely, $\tau = \xi$, and the matrix ;

$$T(\xi, 0) = \begin{bmatrix} \cos\xi & \sin\xi & 0 \\ -\sin\xi & \cos\xi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (7)$$

The total transformation from the reference frame of virtual bond $i + 1$ to that for i is effected by multiplying the vector by the product of these three matrices, or ;

$$T_i = T(\phi_i, \psi_i) = T(\xi, 0) T(\theta, \pi - \phi_i) T(-\eta, -\psi_i) \quad (8)$$

Up to this point only calculations carried out for a chain in one specified conformation have been considered. In order to extend these methods to include the treatment of chains which may exist in a number of conformations, a method for determining the average of transformation matrices over the different possible conformations is needed. The

average of several transformation matrices can be expressed as an integral ;

$$\langle T_i \dots T_{j-1} \rangle = \frac{\int \dots \int (T_i \dots T_{j-1}) \exp \{-E(l)/RT\} d(l)}{\int \dots \int \exp \{-E(l)/RT\} d(l)} \quad (9)$$

where $E(l)$ represents the energy of conformation (l) , R is the gas constant, and T is absolute temperature. The angle brackets denote the statistical mechanical average over all possible conformations of the chain.

A simplification of equation 9 may be obtained if independent rotational potentials are assumed. The simplest way to test this assumption is via detailed study of pertinent molecular models. When this is done for a polypeptide it can be seen that overlaps obtained in the process of varying one of the angles, either ψ or ϕ , are very dependent upon the value of the other angle. Under close scrutiny, however, it can be seen that such interactions encountered with these interdependent pairs of angles are independent of the values for ϕ and ψ for neighboring residues. Since this is true, the conformational energy can be expressed as a sum :

$$E(l) = \sum_{i=1}^{n-1} E(\phi, \psi) \quad (10)$$

The ability to separate the conformational energy, along with the fact that the transformation of a vector from one reference frame to another is a function of only two angles, permits the factorization of the integrals in equation 9 into separate integrals for each bond. This means that equation 9 can be replaced by the product of ratios

such as ;

$$\langle T_i \rangle = \frac{\int T_i \exp \{-E(l_i)/RT\} d(l_i)}{\int \exp \{-E(l_i)/RT\} d(l_i)} \quad (11)$$

From this it can be seen that the average product of several transformation matrices can be expressed as the product of the averages of the individual transformation matrices.

$$\langle T_i \cdots T_{j-1} \rangle = \prod_{h=i}^{j-1} \langle T_h \rangle \quad (12)$$

It should be emphasized that this simplification is not possible for the great majority of macromolecules, and it is most fortunate that such is the case for polypeptides.

An average transformation matrix can be expressed as ;

$$\langle T_i \rangle = T(\xi, 0) \langle T(\theta, \phi_i) T(-\eta, -\psi_i) \rangle \quad (13)$$

The configuration partition function may be written as ;

$$Z = \sum_{\phi} \sum_{\psi} \exp \{-E(\phi, \psi)/RT\} \quad (14)$$

By using equations 13 and 14, $\langle T_i \rangle$ can be evaluated numerically. To accomplish this, the product of the matrices within the angle brackets in equation 13 is computed at intervals of ϕ and ψ . Each product is then multiplied by the appropriate Boltzmann factor, $\exp \{-E(\phi, \psi)/RT\}$, of the conformational energy at the appropriate values of ϕ and ψ . The sum of these terms is then divided by the configuration partition function, Z , for the residue at identical intervals of ϕ and ψ . This quotient is then multiplied by the matrix $T(\xi, 0)$ to give $\langle T_i \rangle$.

The average transformation matrices have been evaluated for several residues. The average transformation matrices used for computational purposes are listed below ;

$$\langle T_c \rangle^{15}(\text{ala}) = \begin{bmatrix} 0.51 & 0.20 & 0.59 \\ -0.046 & -0.61 & 0.21 \\ 0.65 & -0.23 & -0.30 \end{bmatrix} \quad (15)$$

$$\langle T_c \rangle^{15}(\text{gly}) = \begin{bmatrix} 0.36 & -0.077 & 0 \\ -0.092 & -0.037 & 0 \\ 0 & 0 & -0.012 \end{bmatrix} \quad (16)$$

$$\langle T_c \rangle^{19}(\text{pro}) = \begin{bmatrix} 0.593 & 0.316 & -0.001 \\ 0.117 & -0.503 & -0.003 \\ 0.001 & 0.007 & -0.315 \end{bmatrix} \quad (17)$$

$$\langle T_c \rangle^{20}(\text{ala-pro}) = \begin{bmatrix} 0.50 & 0.14 & 0.81 \\ -0.031 & -0.73 & 0.098 \\ 0.67 & -0.084 & -0.32 \end{bmatrix} \quad (18)$$

$$\langle T_c \rangle^{20}(\text{gly-pro}) = \begin{bmatrix} 0.47 & 0.20 & 0 \\ 0.019 & -0.0021 & 0 \\ 0 & 0 & 0.038 \end{bmatrix} \quad (19)$$

$$\langle T_c \rangle^{21}(\text{pro-pro}) = \begin{bmatrix} 0.423 & 0.474 & 0.548 \\ -0.589 & -0.221 & 0.582 \\ 0.640 & -0.667 & 0.186 \end{bmatrix} \quad (20)$$

These matrices can be used in conjunction with other methods as will be shown shortly, to reproduce experimental results for the

appropriate random coil polypeptides. These results are best illustrated by the calculation of three quantities ; the mean squared end to end distance, $\langle r^2 \rangle$, the mean squared radius of gyration, $\langle s^2 \rangle$, and the characteristic ratio, C_r .

The square of the magnitude of the end to end vector may be written as ;

$$r^2 = \underline{r} \cdot \underline{r} = \left(\sum_{i=1}^n \underline{l}_i \right) \cdot \left(\sum_{j=1}^n \underline{l}_j \right) \quad (21)$$

For purposes of illustration, equation 21 may be rewritten as the following array ;

$$r^2 = \begin{array}{ccccccc} \underline{l}_1 \cdot \underline{l}_1 & + & \underline{l}_1 \cdot \underline{l}_2 & + & \underline{l}_1 \cdot \underline{l}_3 & + & \dots \underline{l}_1 \cdot \underline{l}_n \\ \underline{l}_2 \cdot \underline{l}_1 & + & \underline{l}_2 \cdot \underline{l}_2 & + & \underline{l}_2 \cdot \underline{l}_3 & + & \dots \underline{l}_2 \cdot \underline{l}_n \\ \underline{l}_3 \cdot \underline{l}_1 & + & \underline{l}_3 \cdot \underline{l}_2 & + & \underline{l}_3 \cdot \underline{l}_3 & + & \dots \underline{l}_3 \cdot \underline{l}_n \\ \vdots & & \vdots & & \vdots & & \vdots \\ \underline{l}_n \cdot \underline{l}_1 & + & \underline{l}_n \cdot \underline{l}_2 & + & \underline{l}_n \cdot \underline{l}_3 & + & \dots \underline{l}_n \cdot \underline{l}_n \end{array} \quad (22)$$

From this expression it is obvious that r^2 can be rewritten as the sum of two sums. These are the sum of the diagonal elements of equation 22, $(\underline{l}_x \cdot \underline{l}_x)$, and a sum to account for the upper and lower triangular elements. The latter is multiplied in order to account for both the upper and lower triangular elements. The resulting expression for r^2 is

$$r^2 = \sum_{i=1}^n \underline{l}_i^2 + 2 \sum_{i < j} \underline{l}_i \cdot \underline{l}_j \quad (23)$$

This expression is not in a very convenient form for the calculation of r^2 for large chains. In order to make the problem tractible,

advantage is taken of the fact that the dot product of \underline{l}_i and \underline{l}_j may be expressed as $\underline{l}_i^T \times \underline{l}_j$, where \underline{l}_i^T is the transpose of the bond vector \underline{l}_i , and the vector \underline{l}_j is expressed in the coordinate system of \underline{l}_i . With this in mind this product can be rewritten as ;

$$\underline{l}_i \cdot \underline{l}_j = \underline{l}_i^T (\underline{T}_i \cdots \underline{T}_{j-1}) \underline{l}_j \quad (24)$$

By taking this into consideration equation 23 can be rewritten as ;

$$2 \sum_{i < j} \underline{l}_i \cdot \underline{l}_j = 2 \sum_{1 \leq i < j \leq n} \underline{l}_i^T (\underline{T}_i \underline{T}_{i+1} \cdots \underline{T}_{j-1}) \underline{l}_j \quad (25)$$

The square of the magnitude of the end to end vector can now be expressed as ;

$$r^2 = \sum_{i=1}^n \underline{l}_i^2 + 2 \sum_{1 \leq i < j \leq n} \underline{l}_i^T (\underline{T}_i \underline{T}_{i+1} \cdots \underline{T}_{j-1}) \underline{l}_j \quad (26)$$

If the statistical mechanical average is desired then all that is needed is the substitution of the average transformation matrices into equation 26.

Matrix methods have been developed which make the calculation of quantities such as $\langle r^2 \rangle$ relatively easy to accomplish with the aid of a computer. ¹⁸ This is done by making use of what are known as generator matrices. A generator matrix is a matrix, which upon serial multiplication or some other straightforward operation, yield the desired entity. As an example the formulation of such a matrix to calculate $\langle r^2 \rangle$ is outlined below. It can be seen from equation 26 that the calculation of $\langle r^2 \rangle$ requires the summation of three classes of terms when $n > 1$. These three classes of terms can be distinguished as follows ; i) terms for which $n \leq i < j$, ii) terms with $i < n \leq j$, and iii)

terms where $i \leq j < n$. The terms in class one are multiplied by l_i^T at stage i . Class two terms have acquired l_i^T but not l_j at stage i . Lastly, the terms in class three have acquired l_i^T and l_j prior to step i , and are therefore complete. The generator matrix which generates these terms and hence calculates $\langle r^2 \rangle$ is the following 5×5 matrix ;

$$G_i = \begin{bmatrix} 1 & 2l_i^T \langle T_i \rangle & l_i^2 \\ 0 & \langle T_i \rangle & l_i \\ 0 & 0 & 1 \end{bmatrix} \quad (27)$$

The mean squared end to end distance is obtained from the following operation involving G_i ;

$$\langle r^2 \rangle = G_1 G_2 \dots G_n \quad (28)$$

where G_1 and G_n are the first row and the last column of G_i , respectively.

Another quantity of interest which can be calculated in a similar manner is the radius of gyration s , and s^2 . The radius of gyration of a chain molecule is defined as the root mean squared distance of a group of atoms from their common center of gravity. This can be calculated for the alpha carbons of a polypeptide chain using methods similar to those outlined for the calculation of $\langle r^2 \rangle$.

Consider a chain with $n + 1$ alpha carbons joined by n bonds, and if s_i is the distance of alpha carbon i from the center of gravity of the chain for a certain conformation, then ;

$$s^2 = (n + 1)^{-1} \sum_{i=0}^n s_i^2 \quad (29)$$

This can be rewritten in terms of the intrachain distances as follows ;

$$s^2 = (n+1)^{-2} \sum r_{ij}^2 \quad (30)$$

where r_{ij} is the distance between elements i and j . (see reference 10, appendix A) It was shown earlier that the distance between two elements can be written as ;

$$r_{ij}^2 = G_{i+1} (G)^{j-i-2} G_j, \text{ for } j \geq i+2 \quad (31)$$

By substituting this into equation 28, the following expression is obtained :

$$s^2 = (n+1)^{-2} \sum_{i < j-i} G_{i+1} (G)^{j-i-2} G_j + \sum_{j=1}^n l_j^2 \quad (32)$$

The generator matrix required to calculate s^2 , S_i , is closely related to the generator matrix required for the calculation of r^2 . The squared radius of gyration is obtained by the following operation ;

$$s^2 = (n+1)^{-2} \underline{S}_1 (\underline{S}_i)^{n-2} \underline{S}_n \quad (33)$$

where \underline{S}_i is the required generator matrix, and \underline{S}_1 and \underline{S}_n are the first row and the last column of the generator matrix, respectively. The required generator matrix is ;

$$\underline{S}_i = \begin{bmatrix} 1 & 1 & 2\underline{l}_i^T \underline{l}_i & \underline{l}_i^2 & \underline{l}_i^2 \\ 0 & 1 & 2\underline{l}_i^T \underline{l}_i & \underline{l}_i^2 & \underline{l}_i^2 \\ 0 & 0 & \underline{l}_i & \underline{l}_i & \underline{l}_i \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (34)$$

The mean squared radius of gyration can be calculated as shown previously for $\langle r^2 \rangle$, by simply substituting the average transformation matrices into equation 33.

Now that the methodology for determining average transformation matrices, and their uses in calculating the mean squared end to end distance, and the mean squared radius of gyration, the next important question to answer is how these calculated quantities can be related to measured properties of real chain molecules in solution. Before this can be done, however, it is necessary to introduce several concepts from polymer chemistry. Among these are excluded volume effects, the unperturbed state, and the characteristic ratio.

The concept of excluded volume is easy to visualize. When working with large molecules it must be realized that in addition to short range interactions, effects of very long range can also be present. The term long range here refers to interactions between chain segments which are far apart in the actual sequence of the chain. Excluded volume effects are simply a consequence of the fact that the conformations of a chain molecule in which segments of the molecule occupy the same volume of space are not physically allowable. This is a problem which arises as a chain increases in the degree of polymerization, and is a very important consideration at high degrees of polymerization. Another important consideration is the inherent flexibility of the chain. More flexible chains feel this effect at shorter chain lengths than more rigid or "stiffer" chains.

It is intuitively obvious that if a chain experiences excluded volume effects, then the measured dimensions of the chain will be larger than if the chain does not experience this effect. The

calculations outlined thusfar have not taken this effect into account, and hence the quantities calculated in this manner are designated as those for a chain which is unperturbed. This is designated by a subscript ₀, e.g. $\langle r^2 \rangle_0$ denotes the unperturbed mean squared end to end distance. This is related to the mean squared end to end distance, which is an experimentally accessible quantity, by the following relationship ;

$$\langle r^2 \rangle = \alpha^2 \langle r^2 \rangle_0 \quad (35)$$

where α is the expansion factor.²² An important contribution to polymer science was the discovery and description of conditions under which α is equal to unity by P. J. Flory. This is referred to as the theta point, or theta temperature or solvent. Under such conditions α is equal to unity, and therefore $\langle r^2 \rangle = \langle r^2 \rangle_0$. This can be characterized experimentally in a number of ways. One example is that at the theta point van't Hoff's equation for the osmotic pressure for a solution holds over a large range of concentration. This is comparable to the Boyle temperature for a gas, as this is the temperature at which the ideal gas law holds over a wide range of pressures. The important point here is that by using this information, not only is it possible to compare experimental and calculated values, but a mechanism is also available for correcting such quantities as $\langle r^2 \rangle$ and $\langle s^2 \rangle$ for excluded volume effects.

An instructive hypothetical chain molecule is the random flight or freely jointed chain. A freely jointed chain is a linear chain of n bonds of fixed lengths for which the angles at the bond junctions are free to assume any conformation with equal probability. This means

that the direction of neighboring bonds are totally unrelated. It follows from this model that for a freely jointed chain ;

$$\langle \underline{l}_i \cdot \underline{l}_j \rangle = 0, \quad i \neq j \quad (36)$$

Since,

$$\langle r^2 \rangle = nl^2 + \sum_{i < j} \underline{l}_i \cdot \underline{l}_j, \quad (37)$$

it follows that for a freely jointed chain ;

$$\langle r^2 \rangle_0 = nl^2 \quad (38)$$

The characteristic ratio for a given molecule is the ratio of the unperturbed mean squared end to end distance to that for the same molecule assuming that it is freely jointed.

$$C_r = \langle r^2 \rangle_0 / nl^2 \quad (39)$$

This ratio is obviously unity for a freely jointed chain, and the deviation of this ratio from unity for an actual molecule is a measure of the deviation of the molecule from a freely jointed chain. $\langle r^2 \rangle_0$ can be calculated and determined experimentally under certain conditions, and these values when used to calculate the characteristic ratio go a long way in describing the structural properties which determine to a great extent the conformational properties of a chain. This fact is very poignantly illustrated by a comparison of the values for the characteristic ratio for polyglycine, poly-L-alanine, and poly-L-proline. The measured and calculated values for these polymers are 2.16^{15} , 9.27^{15} , and 12.4^{21} , respectively. It can be

said that polyglycine is more flexible than poly-L-alanine, which is in turn more flexible than poly-L-proline. This can be directly related to the local structure of these homopolypeptides. Polyglycine has the greatest degree of rotational freedom, as it has no sidechain. The result for poly-L-alanine is characteristic of most polypeptides for which the characteristic ratio has been determined. Alanine is the simplest residue with a sidechain, and the interactions which arise as a result of the addition of the methyl group limit the conformational space available considerably in comparison to polyglycine. The hindrance to rotation about ϕ is quite severe in poly-L-proline. The value of ϕ is limited to a short range of values centered at about 120° , as the angle ϕ describes a rotation about a bond within a pyrrolidine ring. This means that poly-L-proline is even stiffer than poly-L-alanine, and hence has a higher value for the characteristic ratio.

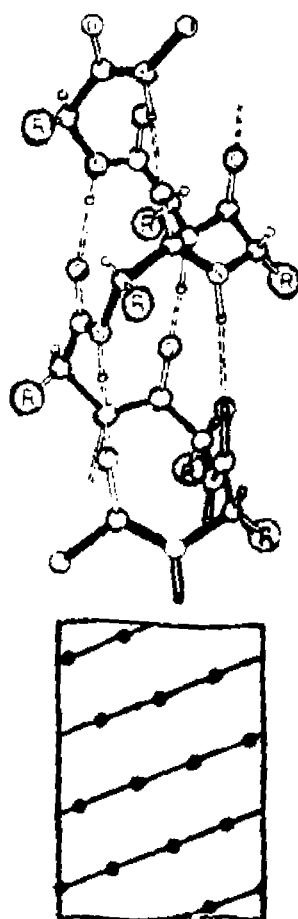
Now that the conformational properties of random coil polypeptides have been discussed, it is of interest now to discuss several specific structures which occur in proteins and synthetic polypeptides. The two most important secondary structures are the alpha helix and beta structure. The discussion here is limited mainly to the alpha helix and only several references will be made to beta structure. The alpha helix is the most abundant secondary structure in proteins. The existence of the alpha helix was first postulated by Linus Pauling in 1951.⁴ The existence of this structure in proteins was soon confirmed from x-ray diffraction studies on crystalline hemoglobin.²³ This was followed by evidence for the existence of the alpha helix in solution.²⁴

The structure of the alpha helix is depicted in figure 5. The alpha helix is also referred to as the 3.6_{13} helix. This arises from the fact that there are 3.6 amino acid residues per turn, and there are 13 atoms within the hydrogen bonded loop. The coordinates for the right-handed alpha helix on a Ramachandran type map are as follows ; $\phi = 133^\circ$, and $\psi = 122.8^\circ$. The supplement of the $N - C^\alpha - C'$ bond is 70.8° .¹⁶ This point is in the center of an allowed region in figure 2. There are several factors which contribute to the stability of this structure. One is the fact that the dipoles of the groups which participate in hydrogen bond formation are aligned. This means that they are in the geometry of minimum energy. Another factor is that the radius of the helix allows for attractive van der Waals interaction across the helical axis. Another consideration is the staggering of the side chains. This can be seen from inspection of figure 5.

It was discovered that a polypeptide could change reversibly from a random coil to alpha helix or vice-versa by changing any number of experimental parameters such as solvent composition.²⁵ The change in conformation or transition is quite striking as the change occurs over a very narrow range of solvent composition. This observation led early investigators to believe that the transition was a cooperative process. The best theoretical framework for treating the helix-coil transition is statistical mechanics. The first such treatment was published by Schellman in 1955.²⁶ His model involved an all or nothing treatment of the transition. Theories which followed allowed for the existence of both helical and coil segments within the same molecule.²⁷⁻³² The work presented in this dissertation is based on the treatment presented by Zimm and Bragg.²⁸

Figure 5. A structural representation of the alpha (3.6₁₃) helix. ³

Shown below this is a cylindrical plot, which gives a good view of the disposition of the sidechains.



In classical statistical mechanics great use is made of what are known as partition functions. A partition function is a rigorous description of a group of atoms or molecules at equilibrium in so far as the partitioning of energy among the various states or levels. The four most frequently used partition functions are the translational, rotational, vibrational, and electronic partition functions. When dealing with statistical mechanics as it is applied to macromolecules, however, another type of partition function is formulated. It is known as the conformational partition function, and it describes the distribution of chain segments in the conformational states of the chain. For a chain of N monomers the conformational partition function is shown rigorously to be ;

$$Z(N) = \int \prod_{i=1}^N \exp(-E_i/kT) dz \quad (40)$$

If it is assumed that each E_i is independent and limited to discrete values, then the above integral can be replaced by the following sum ;

$$Z(N) = \sum_{j=1}^{m^n} \prod_{i=1}^n \exp(-E_i/kT) \quad (41)$$

where m is the number of such discrete values.

The quantity $\exp(-E/kT)$ is known as a statistical weight. This quantity is directly related to the propensity of a given amino acid residue to exist in a particular conformational state. According to the Zimm-Bragg formulation, a residue can exist in one of two possible states. It can exist in the conformation corresponding to the alpha helix, or it can be in one of the many states which correspond to the random coil. The ultimate objective here is to describe the

thermodynamics associated with the transition of a given chain from one conformation to the other. This does not require a rigorous treatment of the segments in a quantum mechanical sense, but it does require the accurate assignment of the statistical weights of the residues relative to that for residues in a coil state.

For illustrative purposes, imagine that a polypeptide chain can be represented by a string of letters which indicate the conformational states of the residues, e.g. ;

... cccccchhhhhhhhhhhccccchhhhhhhhhhhccccchhhhhhhhh ...

where c and h represent residues in either a coil or helical state. For the simplest case of a linear chain with only nearest neighbor interactions, the statistical weight for each residue can be assigned as follows ; i) the quantity u for every c, or coil residue, ii) the quantity s for every h, or helical segment, which propagates a helical segment, and iii) the quantity σs for every h, or helical segment which follows three or more coil residues.

The simplest model for describing the thermodynamics of the transition is a chain in which the Gibbs free energy of a helical segment n units long is seen to vary linearly with n , with a correction for the ends of the helical segments. This can be expressed as follows ;

$$g(n) = g_{\text{ends}} + n \Delta g_{\text{unit}} \quad (\text{for large } n) \quad (42)$$

In this context σ and s can be described in the following terms ;

$$\sigma s = \exp(-g_{\text{ends}}/RT) \quad (43)$$

$$s = \exp(-g_{\text{unit}}/RT) \quad (44)$$

Since

$$K_{eq} = \exp(-G/RT) \quad (45)$$

then σ and s are actually equilibrium constants for helix initiation and propagation, respectively.

The process of initiation is much more difficult than is propagation. This is reflected by the fact that the values of σ are usually several orders of magnitude smaller than the values of s . A helical segment is difficult to start because three pairs of dihedral angles must be fixed in order to form one hydrogen bond, while only one pair must be fixed to propagate a helical segment. The initiation step is difficult from a thermodynamic standpoint as two residues which are fixed in the conformation of the alpha helix have neither the enthalpy of the hydrogen bond nor the entropy of the random coil.

It has been shown previously in this chapter that the partition function can be written as the sum of the statistical weights for each possible conformation. Since two possible states have been assumed for each residue the number of possible conformations for a chain of n residues is equal to 2^n . This means that while the evaluation of Z is relatively simple at small values of n , when n reaches 300 the number of possible conformations is an astronomical 2.037×10^{90} . Listed in table 1 are the eight conformations available to a chain with three residues and the statistical weight for each conformation. The partition function, Z , is the sum of these statistical weights.

Matrix methods have been developed which make the calculation of Z for long chains relatively easy. For purposes of discussion a

TABLE 1.

<u>CONFORMATION</u>	<u>STATISTICAL WEIGHT</u>
ccc	1
cch	σs
chc	σs
hcc	σs
chh	σs^2
hhc	σs^2
hch	$\sigma^2 s^2$
hhh	σs^3
$Z = 1 + 3\sigma s + 2\sigma s^2 + \sigma^2 s^2 + \sigma s^3$	

polypeptide chain with only nearest neighbor interactions will be considered. The statistical weight matrix for such a chain is a 2×2 matrix. In this matrix the state of residue i is described by the columns, while the state of the preceding residue is described by the rows. This statistical weight matrix is ;

$$U_i = \begin{array}{c|cc} & \begin{array}{c} i-1 \\ c \\ h \end{array} & \begin{array}{c} i \\ c \\ h \end{array} \\ \hline \begin{array}{c} c \\ h \end{array} & \begin{array}{cc} 1 & \sigma s \\ 1 & s \end{array} \end{array} \quad (46)$$

This matrix when used in the following operation yields Z ;

$$Z = \underline{J}^* \underline{U}_i^n \underline{J} \quad (47)$$

where \underline{J}^* is row(1,0) and \underline{J} is col(1,1), and n is the number of residues for the homopolymer.

It is of interest now to explore the uses of this partition function. Several quantities can be calculated from the partition function, some of which are experimentally accessible and some which are not, but are of great interest anyway. The main experimentally accessible quantity which can be extracted from the partition function is the fraction of the residues which are in the alpha helical state. Several other quantities which can also be calculated are ; the average number of helical segments, the average length of a helical segment, and the probability that a given residue is in a helical state. The results of this last calculation can be displayed in the form of a helix probability profile, which can be useful in predicting which portion or portions of the chain have the highest

probability for helix formation.

Before the calculation of any of the above quantities is discussed there are several mathematical points which should be discussed. One of these is how some of these quantities can be expressed as a simple derivative of Z . Another important point is how a statistical weight matrix can be diagonalized, and several of these quantities can actually be expressed as analytic expressions for homopolypeptides. Finally the method for calculating these quantities as direct matrix products will be discussed.

The calculation of fraction helix is a good example to use to illustrate how such quantities can be expressed as derivatives of Z . For a chain with n units, Z can be expressed as follows ;

$$Z = 1 + \sum_{j,k} \Omega_{j,k} \sigma^j s^k \quad (48)$$

where Ω_k is the number of ways of combining k helical units into a chain of n residues. This parameter can be shown to be ;

$$\Omega_k = (n - k + 1) \quad (49)$$

If equation 49 is substituted into equation 48, then Z can be evaluated. The terms needed are of the form $\sum_k s^k$ and $\sum_k k s^k$, and the summations are taken from $k=1$ to $k=n$. The first is the following geometric series ;

$$\sum_{k=1}^n s^k = (s^{n+1} - s)/(s - 1) \quad (50)$$

Differentiation of equation 50 yields the other required expression ;

$$s \frac{d}{ds} \sum_{k=1}^n s^k = \sum_{k=1}^n k s^k = \{ s/(s-1)^2 \} \{ n s^{n+1} - (n+1) s^n + 1 \} \quad (51)$$

By substituting equations 49 and 50 into equation 48 Z can be expressed as follows ;

$$Z = 1 + \{ \sigma s^2 / (s-1)^2 \} \{ s^n + n s^{-1} - (n+1) \} \quad (52)$$

The probability $\rho(k)$, that a given chain has k helical units is ;

$$\rho(k) = (n - k + 1) \sigma s^k / Z \quad (53)$$

Therefore,

$$f_h = \sum_{k=1}^n \{ k(n - k + 1) \sigma s^k \} / nZ \quad (54)$$

$$= s (\partial Z / \partial s) / nZ \quad (55)$$

$$= n^{-1} (\partial \ln Z / \partial \ln s)_{n, \sigma} \quad (56)$$

Other quantities can be calculated in a similar fashion. The average number of helical segments is ;

$$n_\sigma = (\partial \ln Z / \partial \ln \sigma)_{n, s} \quad (57)$$

and the average number of residues in a helical segment is ;

$$(\partial \ln Z / \partial \ln s)_{n, \sigma} / (\partial \ln Z / \partial \ln \sigma)_{n, s} \quad (58)$$

A statistical weight matrix which can be diagonalized is a great asset for simplifying the mathematics of a given system. This process is a rather simple one for the simple 2×2 matrix in equation 46. The characteristic equation and roots of the matrix in equation 46 are

$$\begin{bmatrix} 1 - \lambda & \sigma s \\ 1 & s - \lambda \end{bmatrix} = 0, \text{ and} \quad (59)$$

$$\lambda_{1,2} = \frac{(1+s) \pm \{(1-s)^2 + 4 \sigma s\}^{\frac{1}{2}}}{2} \quad (60)$$

where λ_1 and λ_2 are the larger and smaller eigenvalues. The transformation matrix which satisfies the condition ;

$$\underline{T}^{-1} \underline{U}_i \underline{T} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \quad (61)$$

is ;

$$\underline{T} = 1/\lambda_1 - \lambda_2 \begin{bmatrix} 1 & s - \lambda_2 \\ -1 & \lambda_1 - s \end{bmatrix} \quad (62)$$

and it's inverse,

$$\underline{T}^{-1} = \begin{bmatrix} \lambda_1 - s & \lambda_2 - s \\ 1 & 1 \end{bmatrix} \quad (63)$$

The partition function may be written as follows ;

$$Z = J^* \underline{T}^{-1} \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}^n \underline{T} J \quad (64)$$

If $\lambda_1 \neq 1$, and $\lambda_2 \neq 1$, then Z can be written as ;

$$Z = \frac{\lambda_1^{n+1} (1-\lambda_2) + \lambda_2^{n+1} (1-\lambda_1)}{\lambda_1 - \lambda_2} \quad (65)$$

The calculation of fraction helix is used again here for illustrative purposes; this time for the derivation of an analytic expression for the calculation of fraction helix. At reasonable values of σ and s ; σ & $s > 0$ $\lambda_1 > 1$ $\lambda_2 > 0$, as n approaches ∞ ;

$$Z \rightarrow \frac{(1 - \lambda_2) \lambda_1^{n+1}}{\lambda_1 - \lambda_2} \quad (66)$$

As n approaches ∞ , the expression for fraction helix can be shown to converge to the following simple expression ;

$$f_h = \left(\partial \ln \lambda_1 / \partial \ln s \right)_{n,\sigma} = \frac{\lambda_1^{-1}}{\lambda_1 - \lambda_2} \quad (67)$$

Other quantities can also be simplified in a similar manner. It should be emphasized that these expressions are only applicable to chains of very high molecular weight, and only to homopolypeptides.

As noted previously, the value of fraction helix is experimentally accessible, and the preceding expressions have been used to determine the value of σ and s from studies of the appropriate homopolypeptides of varying molecular weight at various temperatures. Unfortunately, most of the naturally occurring amino acids in the form of homopolypeptides do not undergo a helix-coil transition in the temperature range $0^\circ - 100^\circ\text{C}$, or even worse are not soluble in water in this form. H. A. Scheraga and his coworkers have determined the helix-coil stability constants using a regime known as the host-guest technique.^{33,34} They used two residues of the series ω -hydroxyalkyl-L-glutaminy, namely, N^5 - ω -hydroxybutyl-L-glutaminy and ω -hydroxypropyl-L-glutaminy as the host residues, and the amino acid which is

to be characterized is the guest residue. Random copolymers of varying molecular weight and composition are synthesized. The helix initiation and propagation parameters of the guest amino acid residue are then determined by studying the effect of incorporating varying amounts of the guest residue into the polymers. The data is analyzed via several approximate theories which are described in greater detail in chapter 4.³³ Scheraga and his coworkers have determined σ and s for seventeen of the twenty naturally occurring amino acid residues in water.³⁵⁻⁵¹ These values are collected in table 2. Residues for which these parameters have not been determined are estimated by analogy to similar amino acid residues for which σ and s have been determined.

In this last segment two mathematical methods will be discussed which have had a very profound effects in this area of research, and also many varied applications. Both of these methods are due to R. L. Jernigan.⁵¹ The first method involves the calculation of average quantities as a direct matrix product. This makes the calculation of many quantities relatively easy with the aid of a computer. Among these are the probability that a helical state exists at residue i , and the fraction of the residues which are in a helical state. The second method provides a way of combining the information contained in the statistical weight matrix with the powerful computational methods of generator matrices to calculate configuration dependent properties of molecules which are totally or partially helical.

In order to calculate the probability that residue i is in a helical state, $f_{h;i}$, a new matrix U_i^1 must be introduced. This matrix differs from that in equation 46 only by the fact that both elements

TABLE 2

Staistical Weights Used For Amino Acid Residues in Water ^a

<u>Residue</u>	<u>$\sigma \times 10^4$</u>	<u>s at 30 C</u>	<u>Reference</u> ^b
Ala - A	8.0	1.058	36
Asp - D	50.0	0.63	45
Glu, Gln-E,Q	6.0	0.97	41
Phe, His-F,H	18.0	1.069	39
Gly - G	0.1	0.615	35
Ile - I	55.0	1.11	50
Lys - K	1.0	0.947	43
Leu - L	33.0	1.14	38
Met - M	54.0	1.15	47
Asn - N	0.1	0.806	46
Arg - R	0.01	1.017	44
Ser, Cys-S,C	0.8	0.793	37
Thr - T	0.1	0.836	48
Val - V	1.0	0.97	40,51
Trp - W	70.0	1.06	49
Tyr - Y	66.0	0.96	42

a-Pro is assigned s=0 on steric grounds.

b-Reference is for the first amino acid residue listed.

in the first column are set equal to zero. By using this matrix for a given amino acid residue conformations in which this residue exists in a random coil state are ignored. The fraction of the time that this residue is helical is then simply the following quotient ;

$$f_{h;i} = \frac{U^* U_1 U_2 \dots U_{i-1} U'_i U_{i+1} \dots U_n J}{Z} \quad (68)$$

The calculation of fraction helix, f_h , is just as straightforward, but more complicated. In order to calculate f_h the following sum must be evaluated ;

$$f_h = n^{-1} \sum_{i=1}^n f_{h;i} \quad (69)$$

This can be accomplished easily by making use of what are referred to as "supermatrices", or a matrix which is composed of other matrices. The supermatrix which is used in the calculation of f_h is ;

$$\hat{U}_i = \begin{bmatrix} U_i & U'_i \\ 0 & U_i \end{bmatrix} \quad (70)$$

In order to see how this generates the needed sum, consider the following multiplication ;

$$\begin{bmatrix} U_1 & U'_1 \\ 0 & U_1 \end{bmatrix} \begin{bmatrix} U_2 & U'_2 \\ 0 & U_2 \end{bmatrix} = \begin{bmatrix} U_1 U_2 & U'_1 U_2 + U_1 U'_2 \\ 0 & U_1 U_2 \end{bmatrix} \quad (71)$$

It can be seen from this demonstration that the required summation term is accumulating in the upper right hand quadrant of the product. The calculation of fraction helix is achieved by the following

operation ;

$$f_h = n^{-1} z^{-1} \underline{J}^* \hat{U}^n \underline{J} \quad (72)$$

where \underline{J}^* is row(1,0,0,0) and \underline{J} is col(0,0,1,1).

The last point to be covered in this introductory chapter is how a merger of sorts can be executed between the methods for calculating configuration dependent properties as outlined in the first half of the chapter, and the conformational properties as just outlined. As examples the calculation of the mean squared unperturbed end to end distance and the mean squared radius of gyration for a polypeptide which is partially helical are outlined. The mean squared unperturbed end to end distance is calculated as follows ;

$$\langle r^2 \rangle_0 = z^{-1} \underline{F}_i^n \quad (73)$$

where \underline{F}_i is given by ;

$$\underline{F}_i = (\underline{U}_i \times \underline{E}_5) \begin{bmatrix} \underline{G}_c & 0 \\ 0 & \underline{G}_h \end{bmatrix} \quad (74)$$

where \underline{E}_5 is the identity matrix of order 5, and the zeroes represent the appropriate 5 x 5 null matrices. \underline{G}_c is the generator matrix for the calculation of $\langle r^2 \rangle_0$ for a random coil polypeptide, and \underline{G}_h is the same for a completely helical molecule. The latter is constructed along the same lines as the former except that the average transformation matrix is replaced by the one appropriate for an alpha helical residue , i.e. ;

$$\underline{T}_h = \begin{bmatrix} 0.020 & -0.424 & -0.905 \\ -0.425 & 0.816 & -0.391 \\ 0.905 & 0.393 & -0.164 \end{bmatrix} \quad (75)$$

The mean squared unperturbed radius of gyration can be calculated in a similar fashion ;

$$\langle s^2 \rangle_0 = (n+1)^{-2} Z^{-1} (\underline{U}_i \times \underline{E}_7) \begin{bmatrix} \underline{s}_c & 0 \\ 0 & \underline{s}_h \end{bmatrix} \quad (76)$$

This concludes a detailed outline of the statistical mechanical methods used throughout this dissertation. The methods outlined in this chapter form the basis for the more complex calculational methods used in the remainder of the dissertation. This allows the author to avoid a great deal of redundancy.

PART ONE

CONFORMATIONAL STUDIES OF BIOLOGICALLY ACTIVE PEPTIDES

CHAPTER 2

CONFORMATIONAL PROPERTIES OF HUMAN BETA-ENDORPHIN AND BETA-LIPOTROPIN IN WATER AND IN THE PRESENCE OF ANIONIC LIPIDS

INTRODUCTION

Two peptides which exhibit extraordinary biological activities, and are therefore of great biological significance are β -endorphin⁵³ and β -lipotropin.⁵⁴ Both of these molecules are also quite interesting with regard to their conformational properties. While both peptides lack any appreciable amount of ordered structure in water, both contain a significant amount of alpha helix in the presence of anionic lipids⁵⁵ and detergents.⁵⁶ There is evidence that this phenomenon may be at the heart of the biological activity of β -endorphin. Before discussing these properties, however, it would be helpful to know more about these molecules.

The history of β -lipotropin is considerably longer than that of β -endorphin. Interest in the area of pituitary hormones, as related to adipose tissue metabolism, began in 1931 when it was discovered that extracts of anterior pituitary glands produce ketosis and hyperlipemia.⁵⁷ It took many years for the molecule responsible for these effects to be isolated and characterized. Ovine β -lipotropin was isolated and the sequence was determined in 1965.⁵⁸ The sequence of human β -lipotropin was published in 1976.⁵⁹

There is still some debate as to whether β -lipotropin is a precursor to β -endorphin. Speculation began as soon as β -endorphin was discovered, and was found to be identical in sequence to residues 61 - 91 of β -lipotropin. It is also of interest to note that residues 37 - 58 of β -lipotropin are identical in sequence to the hormonal peptide β -melanotropin. Both of these sequences, residues 37 - 58 and

61 - 91, form a part of the β -lipotropin molecule which has been strictly conserved from an evolutionary standpoint, as evidenced in figure 6.

The discovery of the endogenous opioid peptides β -endorphin and the enkephalins ended literally centuries of speculation. Opium and its various components have been used for medicinal purposes for several centuries.⁶⁰ The hypothesis that opiates must bind to specific receptors in nervous tissue in order to produce their effects was proposed in 1959.⁶¹ Goldstein et. al. postulated that such binding should be stereospecific.⁶² They incubated mouse brain homogenates with tritiated levorphanol (a potent opiate agonist), and a one-hundred fold excess of its inactive stereoisomer, dextrorphan. They defined stereospecific binding as that portion of the binding which was blocked by unlabelled levorphanol, but not by dextrorphan. Dextrorphan was not expected to bind to a stereospecific opiate receptor. They found that stereospecific binding of levorphanol accounted for only two percent of the measured binding activity. Three groups independently reported somewhat different findings in 1973.⁶³⁻⁶⁵ All three groups used similar modifications of Goldstein's original experimental protocol. First, they were able to use lower concentrations of radiolabelled ligand, as they used samples which had much higher specific radioactivity. Second, they washed the homogenates with cold buffer in order to remove unbound or loosely bound ligand molecules. These findings, along with those of another group of researchers who found opiate receptors to be present in the central nervous system of all vertebrates,⁶⁷ led investigators to believe that a substance existed which acted as an endogenous ligand

Figure 6. The amino acid sequences of human, ⁵⁹ ovine, ⁵⁸ and porcine ⁶⁶ β -lipotropin. The conserved positions are underlined. The amino acid residues are shown by the one letter code ; Ala=A, D=Asp, E=Glu, F=Phe, G=Gly, H=His, I=Ile, K=Lys, L=Leu, M=Met, N=Asn, P=Pro, Q=Gln, R=Arg, S=Ser, T=Thr, V=Val, W=Trp, and Y=Tyr. ¹¹

	10	20	30	40	50
<u>HUMAN</u>	ELTGQRLRQG	DGFNAGANDG	EGFNALHSL	LADLVAAEKK	DEGFYRMEHF
<u>OVINE</u>	ELTGERLEQA	RG: EAQAESA	AARAELEYGL	VAEAEAAEKK	DSGPYKMEHF
<u>PORCINE</u>	<u>ELAGAFFEPA</u>	<u>RDFEAPAEGA</u>	<u>AARAELEHGL</u>	<u>VAEAQAAEKK</u>	<u>DEGFYKMEHF</u>

	60	70	80	90
	RWGSPPKDKR	YGGFMTSEKS	QTPLVTLFKN	AIKNAYKKG E
	RWGSPPKDKR	YGGFMTSEKS	QTPLVTLFKN	AIKDAHKKG Q
	<u>RWGSPPKDKR</u>	<u>YGGFMTSEKS</u>	<u>QTPLVTLFKN</u>	<u>AIKDAHKKG</u> Q

and activator of the opiate receptor.

In 1975 Hughes reported a substance from mammalian brain extracts which acted similarly to morphine by way of its inhibition of electrically induced smooth muscle contraction.⁶⁸ Also in 1975, two groups reported a morphine-like substance in extracts of calf, rabbit, and rat brain.^{69,70} Terenius and Wahlstrom reported a similar substance which they isolated from human cerebrospinal fluid.⁶⁹ Two pentapeptides, which are responsible for this activity, were isolated from pig brain.⁷¹ The sequences of these peptides were determined to be YGGFM (methionine enkephalin), and YGGFL (leucine enkephalin). They were found in a relative abundance of approximately 3 : 1 YGGFM/YGGFL. The sequence of methionine enkephalin was found to be identical in sequence to residues 61 - 65 of β -lipotropin, as can be seen from figure 6.

While working on the isolation of melanotropins from camel pituitary glands, C. H. Li and his colleagues isolated a novel peptide, which was eventually named β -endorphin.^{72,73} They were unable to isolate β -lipotropin by the usual procedure, but instead they isolated a untriakontapeptide, which was found to be identical in sequence to residues 61 - 91 of β -lipotropin.⁷³ They found that this peptide had little lipolytic activity, but had considerable opiate activity. Two groups isolated similar untriakontapeptides from porcine pituitary glands, which also exhibited considerable opiate activity.⁷⁴⁻⁷⁶ Two groups then independently reported the isolation of human β -endorphin.^{77,78}

A great deal has been learned about β -endorphin since its discovery.⁷⁹ It has been noted that essentially the entire molecule must

be intact in order to produce analgesia.⁸⁰ Some researchers have been seeking ways to enhance the analgesic effect of the peptide, and they have described certain modifications which have met with some success.⁸¹ There is also evidence that lengthening the chain at the carboxyl terminal end enhances the analgesic effect.⁸² The interaction of β -endorphin with its receptor most likely involves electrostatic interactions between cationic groups on the peptide, and anionic groups at the receptor.^{55,83-84} There is evidence that these anionic groups may be provided by cerebroside sulfate.⁸⁵⁻⁸⁷ Such interactions could possibly result in a significant conformational change in β -endorphin.

As mentioned previously, both β -endorphin and β -lipotropin exhibit conformational properties expected for a peptide in the random coil state when in aqueous solution.^{55,56,87} Several authors have tried using some of the popular secondary structure predictive methods,⁸⁸⁻⁹⁰ which are based on measured properties of globular proteins in the crystalline state, to describe the helical content without a great deal of success.^{56,91} This lack of success really comes as no surprise, when consideration is given to the severe limitations of these methods. For example, in a disordered peptide each residue is greatly exposed to the solvent, while the same residue in a globular protein experiences very little exposure to the solvent. When consideration is given to the fact that the tendency of an amino acid residue to form an alpha helix is dependent upon solvent to a great extent, then it is at once obvious why such methods fail in this instance.

A more realistic approach to rationalizing the conformational

properties of these molecules makes use of matrix methods, which have been used with a great deal of success in treating a wide variety of disordered polymers.¹⁸ As noted in the introduction, these methods can be used to assess not only the helical content of these molecules, but their dimensional properties as well. The parameters used for these calculations were obtained from experiments with polypeptides in dilute solution, and from conformational energy calculations. These calculations are therefore based on the properties of the residues in solution and not in the crystalline state, and are subsequently a great deal more successful at describing the conformational properties of these molecules.

Both the absence of alpha helical structure and the dimensional properties of β -endorphin and β -lipotropin in water are accounted for very well by matrix calculations. β -endorphin and β -lipotropin both undergo extensive conformational changes when they interact with various anionic lipids and detergents.^{55,56,92-95} This induced conformation contains a significant amount of alpha helix.^{55,56,94} These results are similar to those obtained when various cationic homopolypeptides interact with similar anionic molecules.⁹⁶⁻⁹⁹ The amount of helicity induced is in good agreement with that expected from studies with these simpler systems. An important aspect of these calculations is that one can predict which sequence or sequences within the molecule has the highest probability of adopting a helical conformation in the presence of the anionic lipid or detergent molecules.

The calculational methods used are a modification¹⁰⁰ of a previously described method which successfully accounts for both the

helical content, ^{100,101} and the dimensional properties ¹⁰⁰⁻¹⁰² of complexes formed by a large number of proteins with sodium dodecyl sulfate. This modification permits the calculation of the fraction helix, and also the propensity of a given residue to initiate, propagate, or terminate a helical segment. Another advance with this modification is the ability to more evenly distribute the end effects of a given helical segment.

CALCULATIONS

Helix probability profiles were calculated essentially as previously described by Mattice et. al.¹⁰⁰ This method is based on the Zimm-Bragg²⁸ treatment of the helix-coil transition in homopolypeptides. A 3×3 statistical weight matrix, U_i , is formulated for amino acid residue i as follows ;

$$\begin{array}{c}
 \begin{array}{cc}
 & i, i+1 \\
 i-1, i & \swarrow \\
 & \text{cc or ch} \quad \quad \quad \text{hc} \quad \quad \quad \text{hh}
 \end{array} \\
 U_i = \begin{array}{cc}
 \text{cc or ch} & \begin{array}{c} 1 \\ 1 \\ 0 \end{array} \\
 \text{hc} & \begin{array}{c} 0 \\ 0 \\ \sigma^{\frac{1}{2}}s \end{array} \\
 \text{ch} & \begin{array}{c} \sigma^{\frac{1}{2}}s \\ 0 \\ s \end{array}
 \end{array} \quad (77)
 \end{array}$$

The rows index the state of amino acid residues i and $i-1$, while the columns index the states of residues i and $i+1$. This differs from the 2×2 matrix in the introduction by the fact that not only are residues i and $i-1$ being considered, but also $i+1$. An h denotes a helical state, while a c denotes a residue in any other conformation. As before the coil state is assigned a statistical weight of unity. Helical residues are assigned a statistical weight of s if they are in the interior of a helical segment. Amino acid residues at either end of a helical segment are assigned a statistical weight of $\sigma^{\frac{1}{2}}s$. In this manner the end effects are evenly distributed between both ends of a particular helical segment. Most of the elements which are assigned a value of zero are nonsense conformations. The exception to this is the second element in the first row, which since it has been

assigned a value of zero, prohibits the occurrence of a helical segment one residue in size. It should be mentioned that the helix probability profiles calculated using the statistical weight matrix in equation 77 are significantly different from those obtained when the burden of helix initiation is placed on the amino acid residue of the helical segment.¹⁰⁰

The statistical weights for the amino acid residues which were used in the calculations are listed in table 2. The prolyl residue was assigned a value of zero for s in recognition of the severe steric problems encountered upon its incorporation into an alpha helix.¹⁰³ Parameters for the cysteinyl residue were not required because this residue is not present in either β -endorphin^{73,104} or β -lipotropin.^{58,59}

The statistical weights used for the residues in the presence of anionic lipids and detergents are based on a model described previously.¹⁰¹ This model was based on the fact that cationic,^{96-99,105,106} but not anionic¹⁰⁷ or uncharged,^{108,109} homopolypeptides have a marked increase in their tendency for helix formation in the presence of sodium dodecyl sulfate. As a consequence of this fact, larger values of σ and s are used for arginyl, histidyl, and lysyl residues. This model has been shown to be able to account for both the optical activity and transport properties of a large number of sodium dodecyl sulfate-protein complexes.^{101,102} The values used for these residues are ; $\sigma = 0.05$, and $s = 1.7$.

If there is compensation between long range intramolecular interactions and solvent-macromolecule interactions, the configuration partition function, Z , for a partially helical copolypeptide

containing n amino acid residues is ;

$$Z = \underline{J}^* \underline{U}_1^n \underline{J} \quad (78)$$

where \underline{J}^* is row(1,0,0) and \underline{J} is col(1,1,0). The probability that amino acid residue i is in a helical state, $p_{h;i}$, is computed by ;

$$p_{h;i} = Z^{-1} \underline{J}^* \underline{U}_1 \underline{U}_2 \cdots \underline{U}_{i-1} \underline{U}_{h;i}^1 \underline{U}_{i+1} \cdots \underline{U}_n \underline{J} \quad (79)$$

where $\underline{U}_{h;i}^1$ is obtained from \underline{U}_i by nulling all elements in the first column. The overall helical content of a copolyptide is ;

$$f_h = n^{-1} \sum_{i=1}^n p_{h;i} \quad (80)$$

Each $p_{h;i}$ is the sum of the probabilities that an amino acid residue initiates, propagates, or terminates a helical segment. These probabilities are denoted by $p_{b;i}$, $p_{p;i}$, and $p_{t;i}$, respectively. They are computed in the following manner ;

$$p_{b;i} = Z^{-1} \underline{J}^* \underline{U}_1 \underline{U}_2 \cdots \underline{U}_{i-1} \underline{U}_{b;i}^1 \underline{U}_{i+1} \cdots \underline{U}_n \underline{J} \quad (81)$$

$$p_{p;i} = Z^{-1} \underline{J}^* \underline{U}_1 \underline{U}_2 \cdots \underline{U}_{i-1} \underline{U}_{p;i}^1 \underline{U}_{i+1} \cdots \underline{U}_n \underline{J} \quad (82)$$

$$p_{t;i} = Z^{-1} \underline{J}^* \underline{U}_1 \underline{U}_2 \cdots \underline{U}_{i-1} \underline{U}_{t;i}^1 \underline{U}_{i+1} \cdots \underline{U}_n \underline{J} \quad (83)$$

The primed matrices in equations 81 - 83 are obtained by nulling all elements of \underline{U}_i except one. For $\underline{U}_{b;i}^1$ the nonzero element is 1,3, for $\underline{U}_{p;i}^1$ it is 3,3, and it is the 3,2 element for $\underline{U}_{t;i}^1$.

The mean squared unperturbed radius of gyration was calculated in a manner similar to that shown in equation 76. The transformation matrices used are those listed in the introduction. All of these

calculations are straightforward applications of the principles outlined in the introduction.

Pages 58 and 59 are missing in number only.

William J. Cooper, Jr.

William J. Cooper, Jr.

Dean

RESULTS AND DISCUSSION

Beta-Endorphin and Beta-Lipotropin in Water

Extremely small helical contents are computed for human β -endorphin and ovine β -lipotropin in water. At 30°C, these results for f_h are 0.014 for β -endorphin and 0.031 for β -lipotropin. The circular dichroism spectra of these proteins in water indicate the presence of very little, if any, ordered structure. ^{55,56,87} Thus the calculation and the experimental results are in agreement.

The computation of the unperturbed dimensions yields a value of 21 Angstroms for $\langle s^2 \rangle_0^{1/2}$ for β -endorphin and 35 Angstroms for β -lipotropin. It would be of great interest to compare these calculated values with those obtained from experiment. However, only the perturbed dimensions can be determined from the available viscosity data. ⁵⁶ The intrinsic viscosity $\{\eta\}_0$ can be converted to the root mean squared radius of gyration by the following relationship ;

$$\langle s^2 \rangle_0^{1/2} = (1/6)^{1/2} (\{\eta\}_0 M / \phi)^{1/3} \quad (84)$$

where M is the molecular weight, and ϕ is a constant whose value is about 0.0021 for $\{\eta\}_0$ in dl g⁻¹ and $\langle s^2 \rangle_0^{1/2}$ in Angstroms. The resulting values of $\langle s^2 \rangle_0^{1/2}$ are 22 Angstroms for β -endorphin and 36 Angstroms for β -lipotropin. The agreement between calculation and experiment here depends upon how large of a correction must be made to account for excluded volume effects. If the correction is small, as it is for myelin basic protein, ¹¹⁰ then the agreement between the computed and the experimental values for $\langle s^2 \rangle_0^{1/2}$ for β -endorphin and β -lipotropin

is quite good.

Overall Helical Content for β -Endorphin and β -Lipotropin in SDS

Computation using dodecyl sulfate parameters at 30°C yields a value of 0.19 for f_h for β -endorphin and 0.25 for β -lipotropin. In conjunction with figure 4 of reference 101, the predicted value of the mean residue ellipticity at 222 nanometers is $-8200 \pm 1100 \text{ deg cm}^2 \text{ dmol}^{-1}$ for these proteins in dodecyl sulfate. Experimental measurements confirm that a significant amount of helix is induced in these proteins by dodecyl sulfate. However, the measured ellipticities at 222 nanometers are somewhat more negative, being $-12,000$ for β -endorphin and $-14,000$ for β -lipotropin. ⁵⁶

While phosphatidyl choline and cerebroside, which are neutral lipids, have little effect on the circular dichroism of β -endorphin, anionic lipids induce an appreciable amount of helical structure in this protein. ⁵⁵ Circular dichroism spectra show mean residue ellipticities at 222 nanometers of about $-8000 \text{ deg cm}^2 \text{ dmol}^{-1}$ in the presence of cerebroside sulfate or ganglioside G_{M1} , while $-10,000$ is obtained in the presence of phosphatidic acid or phosphatidyl serine. The behavior in the presence of cerebroside sulfate is of great interest since this lipid has been suggested to be an active component of the opiate receptor in the brain. ⁸⁴⁻⁸⁶ The ellipticity observed in the presence of cerebroside sulfate is in excellent agreement with that computed for β -endorphin using the dodecyl sulfate parameters.

Helix Propagation Probability Profile for β -Endorphin in Acidic Lipids

An average length of six residues is computed for the helical

segments of β -endorphin when the dodecyl sulfate parameters are used. Attention is now focused upon the profile constructed from $p_{p;i}$, rather than $p_{h;i}$ in order to suppress contributions from extremely short helical segments. The helix propagation probability profiles for β -endorphin in water and dodecyl sulfate are shown in figure 7. The first five residues of β -endorphin, which have the same amino acid sequence as methionine enkephalin, are seen to have small $p_{p;i}$ both in water and dodecyl sulfate. As a matter of fact it is obvious from figure 7 that the entire amino terminal portion of the chain seems to have very little tendency for alpha helix formation in either solvent. Large values for $p_{p;i}$ are encountered only in the carboxyl portion of the β -endorphin chain, and these values are not obtained unless the dodecyl sulfate parameters are used for the four lysyl residues. The probable helix forming potential of this portion of the molecule has been mentioned previously, ^{55,86} but unambiguous predictions in quantitative agreement with experiment have not been forthcoming.

The area of highest helix propagation probability, residues 17-29 of human β -endorphin, has the sequence LFKNALIKNAYKK. The only side-chains bearing charges at physiological pH are the four lysyl residues. Furthermore, several of the side chains in this sequence are hydrophobic. The disposition of the side chains in this sequence when the backbone forms an alpha helix is shown in figure 8. The helix is seen to have a surface where the side chains are those of leucyl, alanyl, phenylalanyl, isoleucyl, and tyrosyl residues. This makes this surface very strongly hydrophobic. Three cationic residues are found at the carboxyl terminal end of the helical segment, and a fourth occurs at the other end. Thus the region of highest helix

Figure 7. Helix propagation probability profiles for human β -endorphin using water (dotted line) and dodecyl sulfate (solid line) parameters. ¹¹¹

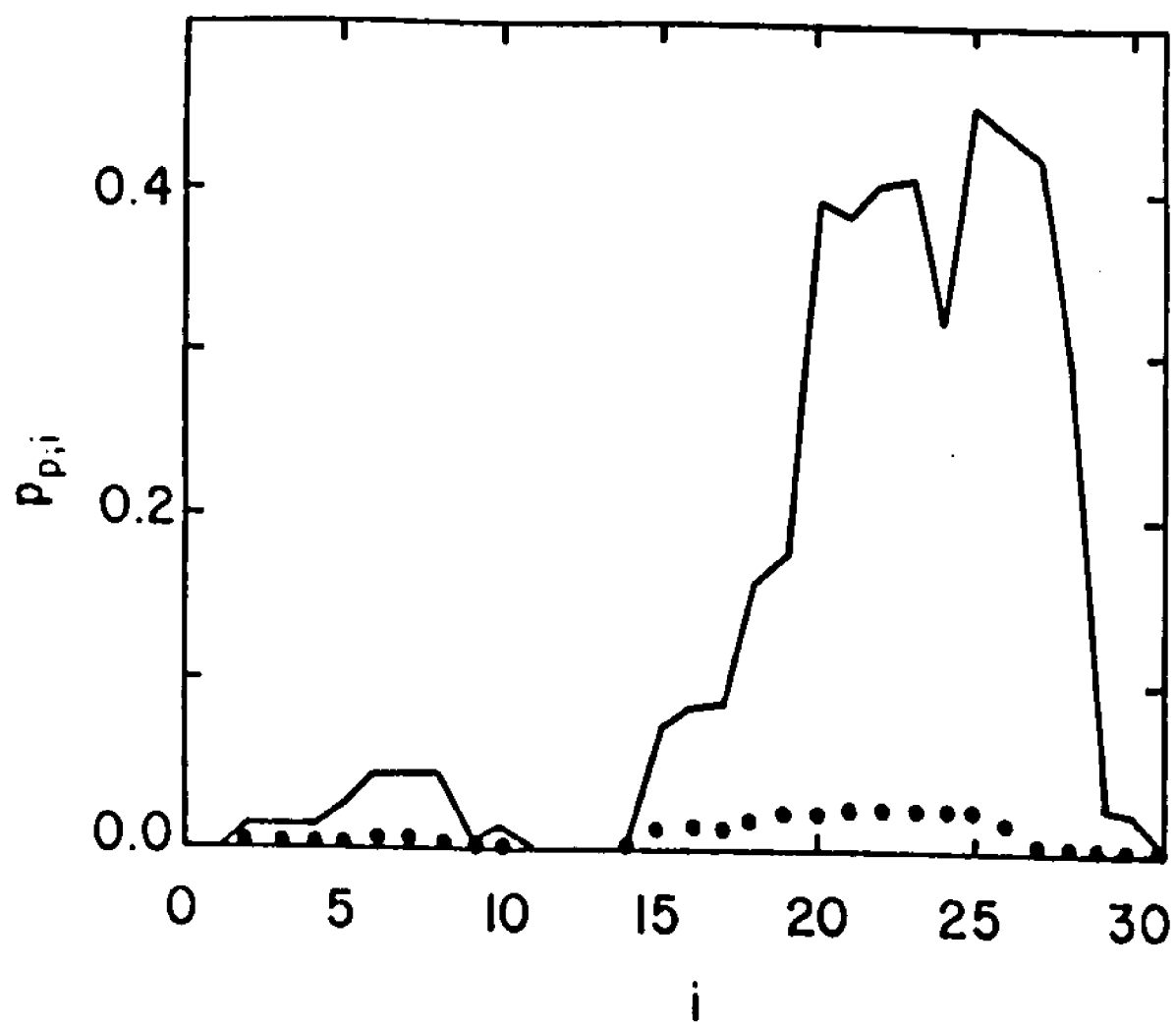
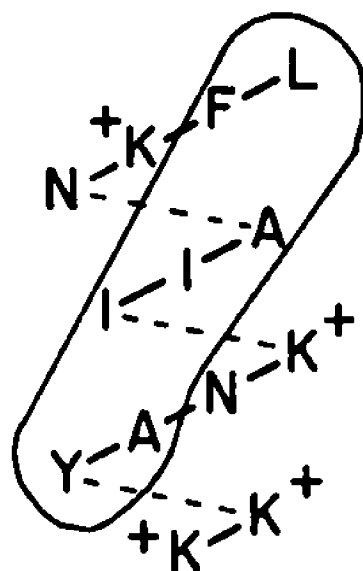


Figure 8. Amino acid residues 17-29 (LFKNALIKNAYKK) for human δ -endorphin arranged to reflect the disposition of the side-chains in the alpha helical conformation. ¹¹¹



RESIDUES 17- 29

forming potential in β -endorphin is ideally suited to participate in favorable hydrophobic and electrostatic interactions with acidic lipids. These results would lend support to a model in which the latter half of the molecule is partially immersed in an anionic membrane when the peptide is bound at the receptor site.

It is of further interest to inquire how this prediction fits with the large amount of work which has been done with β -endorphin, and β -endorphin analogues and fragments. Camel β -endorphin⁷⁸ has the same sequence as human β -endorphin, except for two positions. The differences are at position 27 (camel:H, human:Y), and at position 31 (camel:Q, Human:E). Camel β -endorphin has a higher binding activity than human β -endorphin.^{112,113} However, the analgesic activity is about the same for both peptides. This implies several things. First, the binding activity of opioid peptides is not a satisfactory indication of the analgesic potency of a given molecule. It also points toward the area of the chain which the calculation predicts as being important in the interaction between β -endorphin and cerebroside sulfate as being just that. The substitution of the histidine into position 27 adds yet another cationic sidechain which can interact with the negatively charged receptor, while the replacement of glutamate at position 31 by glutamine eliminates a negatively charged sidechain, which could possibly be a source of unfavorable interaction with a negatively charged receptor.

A study of a localized β -endorphin receptor from rat brain also yielded some interesting results.⁸³ They found that among other things that reasonable concentrations (submillimolar) of divalent cations inhibited the binding activity of β -endorphin. This is

pertinent to the discussion because calcium ions inhibit the formation of alpha helix by β -endorphin in the presence of cerebroside sulfate.⁵⁶ The physiological significance of this fact is not known at the present time. From inhibition patterns with several related peptides it appears that β -endorphin and methionine enkephalin interact with overlapping but nonidentical binding sites.

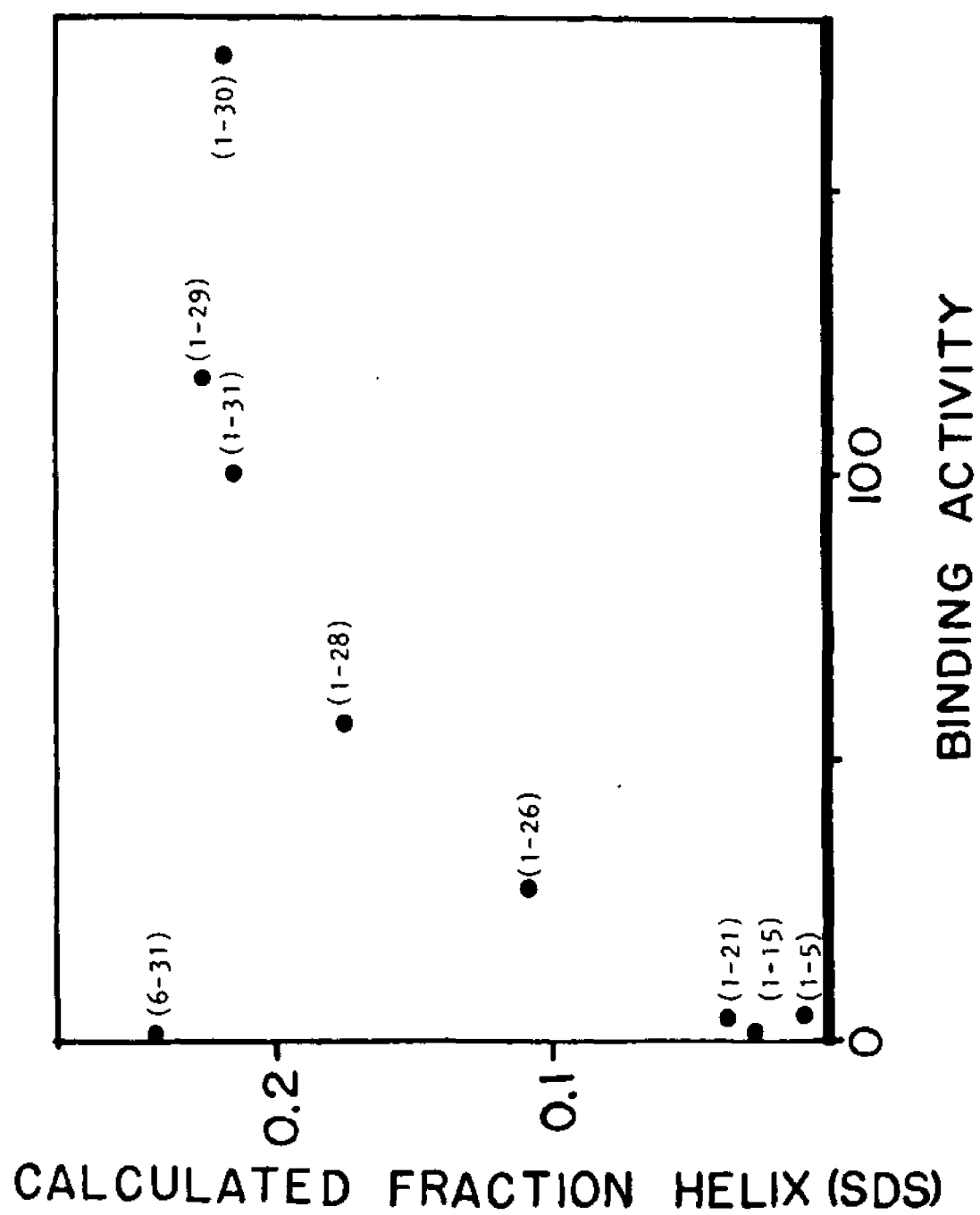
This study led to a detailed study of the binding activity and analgesic properties of a group of β -endorphin fragments of different chain lengths.¹¹⁴ The fragments studied were β -endorphin (1-31), (1-30), (1-29), (1-28), (1-26), (1-21), (1-15), (1-5), and camel β -endorphin (6-31). Both the (1-30) and (1-29) fragments have a higher binding activity than β -endorphin itself. On the other hand, chains which only contain portions of the amino terminal region of the chain have a very small amount of binding activity relative to the intact β -endorphin molecule. The binding activity increases fivefold from (1-15) to (1-21), and sixfold from (1-21) to (1-26). Further additions to the chain further increase the binding activity, but not in near as dramatic fashion as from (1-15) to (1-26). This data also implies that the carboxyl terminal region, and the residues of the postulated helical segment in particular, are important for the binding activity of β -endorphin.

A plot of the binding activity of the peptides versus the calculated f_h for the peptides in dodecyl sulfate is shown in figure 9. The change in the binding activity as the chain length is increased has a strong correlation to the calculated f_h in dodecyl sulfate, with one exception. The exception is camel β -endorphin (6-31). The analgesic potency also increases in a manner roughly proportional to the binding

activity. The peptide camel β -endorphin (6-31) has relatively little binding affinity and little or no analgesic potency. This means that both properties are dependent upon both the amino terminal and carboxyl terminal portions of the chain. The amino terminal region seems to be important in so far as the analgesic potency of the molecule is concerned, but the binding process involves both the methionine enkephalin sequence, residues 1-5, and a portion of the molecule which is towards the carboxyl terminus.

The best explanation is that the amino terminus probably binds to the methionine enkephalin site, and the carboxyl terminal sequence binds, and in some way reinforces the affinity of the β -endorphin molecule for the receptor. It is not known at the present time whether β -endorphin and methionine enkephalin bind differently to the same receptor, or to different receptors with overlapping specificity.

Figure 9. A plot of the binding activity of several β -endorphin fragments relative to human β -endorphin, where the activity of human β -endorphin is assigned a value of 100, versus the value of f_h calculated for the peptides with the dodecyl sulfate parameters. The binding data is from reference 114.



CHAPTER 3

LIPID INDUCED CONFORMATIONAL CHANGES IN GLUCAGON, SECRETIN, AND VASOACTIVE INTESTINAL PEPTIDE

INTRODUCTION

The gastrointestinal hormones glucagon,¹¹⁵ secretin,^{116,117} and vasoactive intestinal peptide^{118,119} have amino acid sequences (figure 10) which suggest that they are derived from a common ancestor.¹²⁰⁻¹²² These similarities are not confined to their covalent structures. While each peptide displays certain unique actions, they also have many biological activities in common.¹²³⁻¹²⁶ One area of great current interest is their potential involvement as neurotransmitters ; especially vasoactive intestinal peptide.^{127,128}

None of these peptides is highly ordered when in dilute aqueous solution. Presumably order is induced when they interact with their receptor. One objective of this chapter is to determine the portion of the chain in which anionic lipids are likely to induce order. The approach used¹⁰⁰ is essentially the same as that used in the previous chapter for β -endorphin. The configuration partition function for a particular peptide is formulated in a manner which quantitatively incorporates conformational effects observed with synthetic polypeptides and proteins in solution.^{100,101} Experimentally measurable optical properties can be extracted from the configuration partition function. If the agreement between computed and experimental values is good, additional information, not necessarily susceptible to direct experimental determination, may legitimately be extracted from the configuration partition function. The important information here is that portion of the identity of that portion of the chain in which order is induced.

Figure 10. Amino acid sequences of glucagon, secretin, and vasoactive intestinal peptide. The carboxyl terminus is an amide group in secretin and vasoactive intestinal peptide.

GLUCAGON

H S Q G T - F T S D Y - S K Y L D - S R R A Q - D F V Q W - L M D T

SECRETIN

H S D G T - F T S E L - S R L R D - S A R L Q - R L L Q G - L V

VASOACTIVE INTESTINAL PEPTIDE

H S D A V - F T D N Y - T R L R K - Q M A V K - K Y L N S - I L N

The study in this chapter is also of interest from a purely theoretical point of view. While the amino acid sequences of glucagon, secretin, and vasoactive intestinal peptide exhibit extensive homology, those differences which do occur are responsible for the prediction that these peptides should show significantly different helical content in the presence of anionic lipids. If the configuration partition function is formulated in the manner used for other peptides and proteins,^{100,101,129,130,111} the helical content in the presence of an anionic lipid is predicted to increase in the order glucagon < secretin < vasoactive intestinal peptide. Thus the behavior of these peptides provides a sensitive test for the validity of assumptions used in the treatment of processes by which anionic lipids induce order in polypeptides and proteins.

EXPERIMENTAL

Materials were purchased from the following sources : vasoactive intestinal peptide from Vega Biochemicals, secretin from United States Biochemical Corporation, glucagon, L- α - phosphatidic acid (dipalmitoyl) and L- α - lysophosphatidylcholine (palmitoyl) from Sigma Chemical Company, phosphatidyl glycerol from Calbiochem, and sodium dodecyl sulfate from Matheson, Coleman, and Bell Manufacturing Chemists.

Circular dichroism measurements were conducted on protein solutions prepared by dilution of a stock solution with distilled deionized water. The pH was measured with a Beckman 3500 pH meter. Circular dichroism measurements were made with a Durrum-Jasco spectropolarimeter calibrated with d-10-camphorsulfonic acid in water.¹³¹ Measurements were made using fused quartz cells with a pathlength of 10 mm. Protein concentrations were about 0.005 mg/ml. Typical uncertainties for $[\theta]_{222}$ were about 300-700 deg cm² dmol⁻¹. Cell temperature was maintained with a circulating water bath connected to a brass water jacket surrounding the cell. Sample temperature was determined with a Yellow Springs Instrument model Y25c telethermometer before and after each spectrum was recorded.

Concentrations of secretin and vasoactive intestinal peptide stock solutions were determined from a quantitative amino acid analysis of the acid hydrolysate. The concentration of the glucagon stock solution was determined from the absorbance at 270 nanometers.¹³²

CALCULATIONS

Helix probability profiles were calculated as described in chapter 2 for β -endorphin.^{100,129} This method is based on the Zimm-Bragg analysis of the helix-coil transition in polypeptides.²⁸ In the method used here, the burden of helix initiation is placed equally on amino acid residues at each end of the helical segment. The values of σ and s used in the calculations are presented in table 2. The probability that a residue is helical, and the probability that a residue initiates, propagates, or terminates a helical segment were calculated in the manner specified by equations 79, and 81 - 83, respectively. These equations represent straightforward applications of the principles outlined in the introduction.

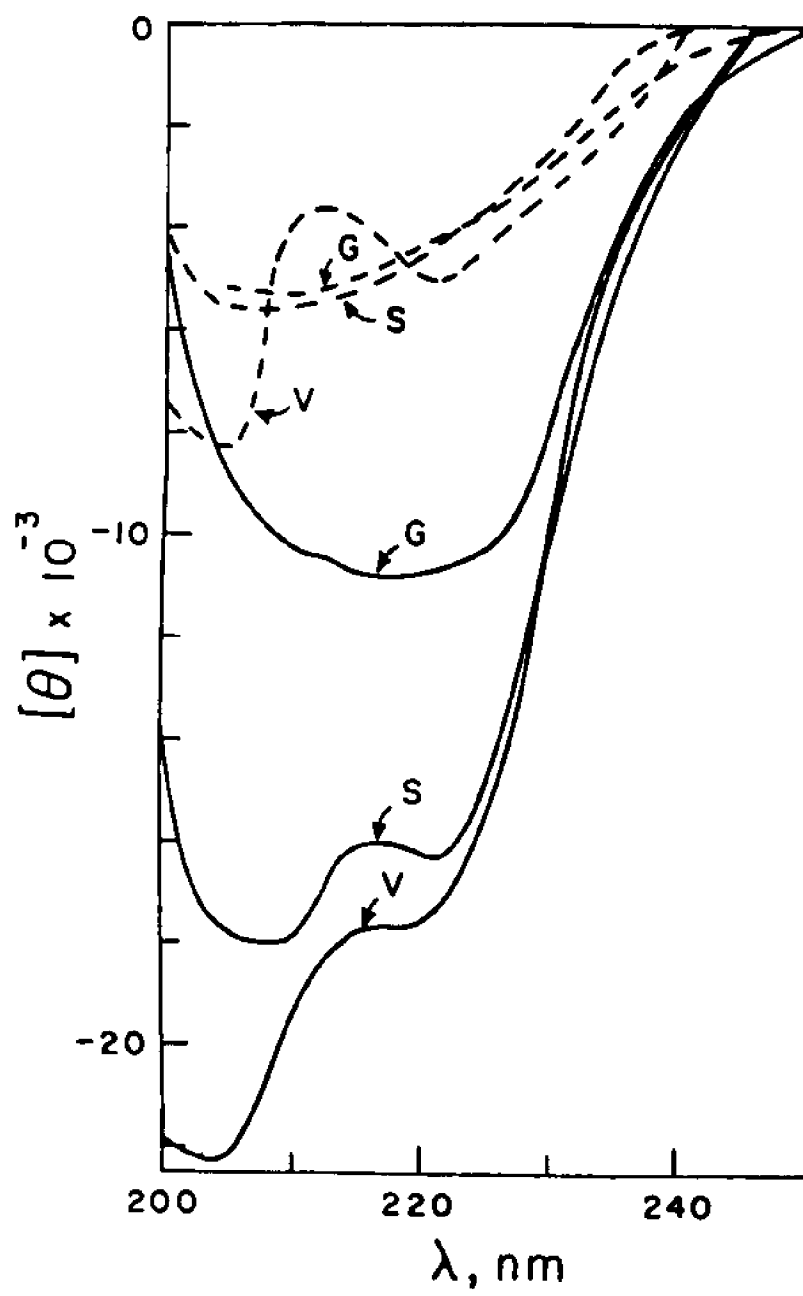
RESULTS AND DISCUSSION

Conformation in Water

The low intensities observed in the circular dichroism of aqueous solutions of glucagon, secretin, and vasoactive intestinal peptide (figure 11) suggest a predominately disordered structure. Temperature coefficients, $d[\theta]_{222}/dT$, are positive and in the range of 40 - 100 $\text{deg cm}^2 \text{ dmol}^{-1} \text{ K}^{-1}$, being the largest for secretin and the smallest for glucagon. In contrast, negative temperature coefficients are obtained for completely disordered homopolypeptides bearing $-\text{CH}_2\text{R}$ sidechains in the L-configuration.^{133,134} Thus the sign of the temperature coefficients seen with glucagon, secretin, and vasoactive intestinal peptide suggests the presence of a limited amount of thermally labile ordered structure. The helical content for glucagon, secretin, and vasoactive intestinal peptide computed using the values of σ and s in table 2 is in the range 4 - 6 %.

Several earlier studies of the conformational properties of these peptides in dilute aqueous solution also reached the conclusion that they contain a small amount of ordered structure. The optical rotatory dispersion of secretin^{135,136} and vasoactive intestinal peptide¹³⁷ in water has been interpreted as signifying the presence of a short helical segment. Glucagon presents a more complicated situation. In sufficiently dilute aqueous solution, the helical content has been found to be similar to that of secretin and vasoactive intestinal peptide.¹³⁸⁻¹⁴² However, glucagon readily undergoes aggregation accompanied by an increase in

Figure 11. Circular dichroism spectra at 30°C for glucagon, secretin, and vasoactive intestinal peptide in water (dashed lines), and in 0.003 M dodecyl sulfate (solid lines). Spectra are denoted G (glucagon), S (secretin), and V (vasoactive intestinal peptide). ¹⁴³



helicity. 137,140,141,144 In the crystalline state glucagon adopts a highly helical conformation which is stabilized by intermolecular interactions. 145

Conformation in Anionic Detergents and Lipids

Figure 11 also depicts circular dichroism spectra observed with glucagon, secretin, and vasoactive intestinal peptide in 0.003 M sodium dodecyl sulfate at 30°C. All three peptides undergo drastic change in the presence of the anionic detergent. The spectra exhibit features expected for proteins which contain an appreciable amount of alpha helix. Glucagon shows the least amount of induced ordered structure. The observed mean residue ellipticity at 222 nanometers ($-10,600 \text{ deg cm}^2 \text{ dmol}^{-1}$ at 30°C) is in harmony with the result reported recently by Wu and Yang. 146 A larger helical content is induced in secretin and vasoactive intestinal peptide. Mean residue ellipticities at 222 nanometers are -16,100 and -17,200, respectively. For all three peptides $[\theta]_{222}$ becomes less negative as temperature increases over the range 7 - 50°C. Temperature coefficients range from $100 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ K}^{-1}$ for secretin to $190 \text{ deg cm}^2 \text{ dmol}^{-1} \text{ K}^{-1}$ for vasoactive intestinal peptide. Thus temperature coefficients seen with glucagon and vasoactive intestinal peptide, but not secretin, become substantially larger in the presence of 0.003 M dodecyl sulfate. The large temperature coefficient for vasoactive intestinal peptide causes it's circular dichroism to differ from that for secretin by a greater amount at 7° than is the case at 30°. At the lower temperature, $[\theta]_{222}$ is $-21,600 \text{ deg cm}^2 \text{ dmol}^{-1}$ for vasoactive intestinal peptide, but only -18,900 for secretin. All three peptides remain

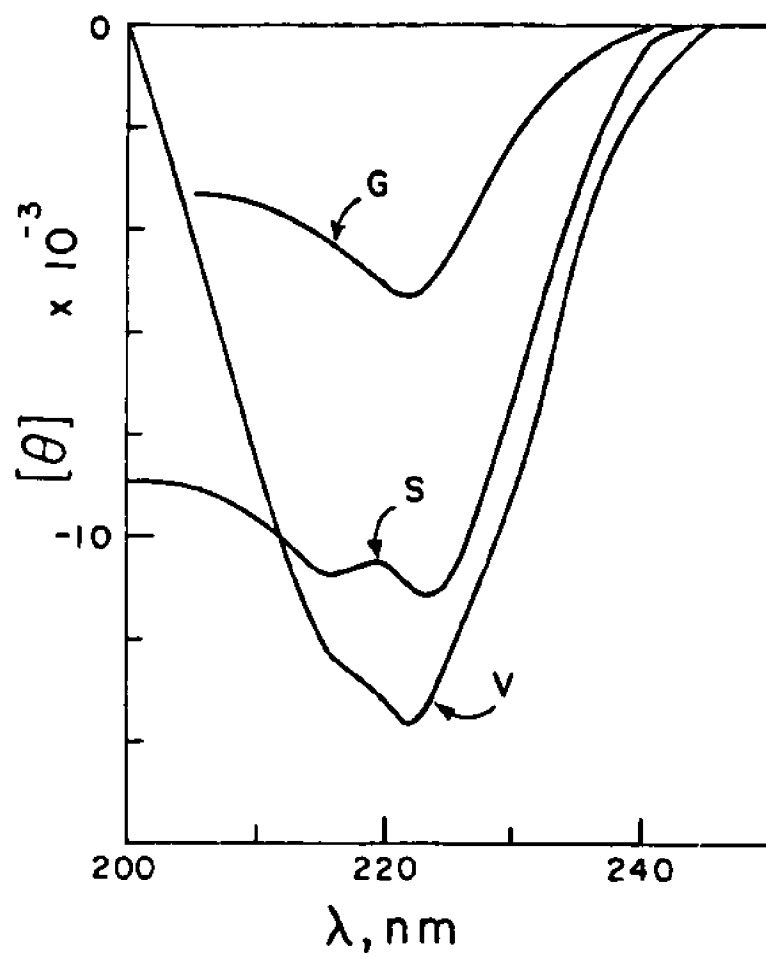
somewhat helical, even at 50°C.

All three peptides also undergo extensive conformational changes in the presence of anionic lipids. Spectra seen in the presence of phosphatidic acid (figure 12) are similar in nature to those obtained in dodecyl sulfate, indicating a significant amount of induced alpha helix. Relative values of $-\left[\theta\right]_{222}$ were the same as in sodium dodecyl sulfate, i.e., vasoactive intestinal peptide > secretin > glucagon. However, for each peptide $\left[\theta\right]_{222}$ is somewhat less negative than was the case in sodium dodecyl sulfate. At 30°C $\left[\theta\right]_{222}$ is found to be -5000, -11,000, and -13,700 deg cm² dmol⁻¹ for glucagon, secretin, and vasoactive intestinal peptide, respectively. The induced helical structure is destroyed in glucagon at 30°C, and in secretin at 50°C. Vasoactive intestinal peptide shows a much smaller dependence for $\left[\theta\right]_{222}$ than do either glucagon or secretin.

Considerable helical content is also induced in these three peptides by phosphatidyl glycerol. At 30°C $\left[\theta\right]_{222}$ is found to be -7000 and -17,300 for glucagon and vasoactive intestinal peptide, respectively. The circular dichroism of secretin in phosphatidyl glycerol was not reproducible. While all measurements showed the induction of significant helicity, it could not be convincingly established whether the resulting helical content was greater or less than that seen with vasoactive intestinal peptide.

Circular dichroism spectra can be converted to helical content using figure 4 of reference 101. This figure depicts measured $\left[\theta\right]_{222}$ for several proteins in sodium dodecyl sulfate as a function of their helical content computed from their amino acid sequences via matrix methods. In this manner the helical content in sodium dodecyl sulfate,

Figure 12. Circular dichroism spectra at 30°C in 4.5×10^{-5} M L- α -phosphatidic acid, dipalmitoyl. Curves are denoted by G (glucagon), S (secretin), and V (vasoactive intestinal peptide). ¹⁴³



phosphatidic acid, and phosphatidyl glycerol is found to be in the range 0.14-0.28 for glucagon, 0.29-0.43 for secretin, and 0.37-0.47 for vasoactive intestinal peptide. Theoretical values were computed using the matrix method in which the interaction with anionic lipids causes an increase in the probability for helix formation by arginyl, histidyl, and lysyl residues. The calculation yields helical contents at the lower end of the range defined by experimental circular dichroism measurements. These values are 0.20, 0.28, and 0.37 for glucagon, secretin, and vasoactive intestinal peptide, respectively.

A proper appreciation for the achievement of the calculation is obtained by recalling the following facts : First, the computation predicts significant helical content is induced in all three peptides by anionic lipids or detergents. Second, the calculation correctly predicts the order of the helical content in the presence of dodecyl sulfate or phosphatidic acid is glucagon < secretin < vasoactive intestinal peptide. Third, the calculation correctly predicts that the helical content induced in vasoactive intestinal peptide is about twice that for glucagon. Finally, these results are obtained without manipulation of any adjustable parameters whatsoever. There has been no "fine tuning" of the model in order to maximize agreement between computation and experiment.

Helix Propagation Probability Profiles

Since it has now been established that the formulation of the configuration partition function successfully captures the essential features of the conformational change induced by anionic lipids and detergents, it is then possible to extract from that configuration

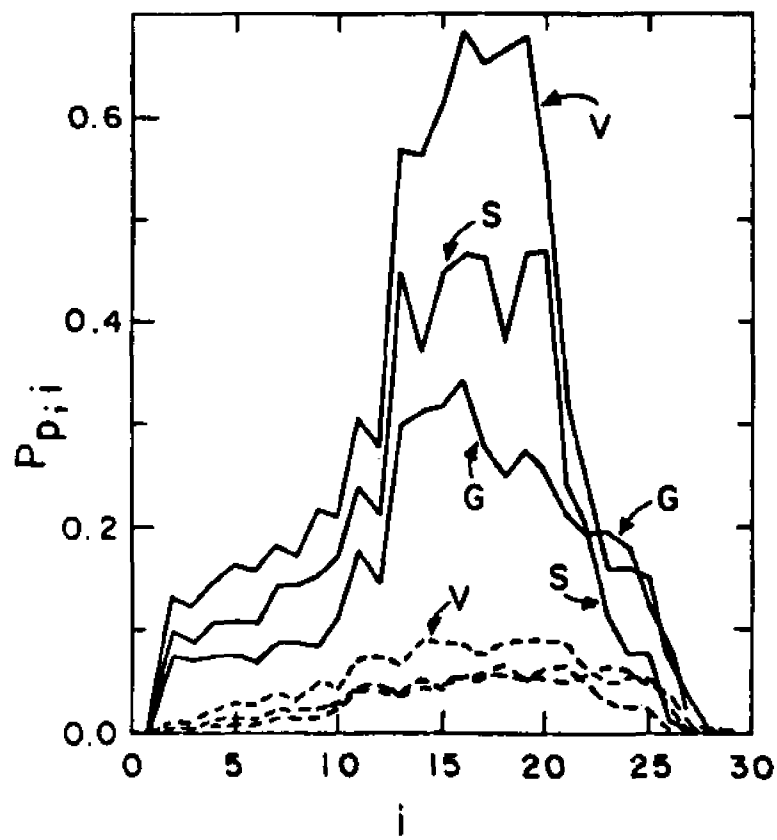
partition function properties which are not readily susceptible to experimental measurement.

$p_{h;i}$ denotes the probability that amino acid residue i is in a helical state. This probability is the sum of the probabilities that amino acid residue i initiates, propagates, or terminates a helical segment. The probability for propagation is denoted by $p_{p;i}$. Profiles depicting $p_{h;i}$ or $p_{p;i}$ as a function of i will differ little from one another if helical segments are long, because there are few ends. However, these two profiles show significant differences if the helical segments tend to be short. The profile based on $p_{p;i}$ partially suppresses the influence of short helical segments. It is this profile for these peptides which is discussed here.

Glucagon

Helix propagation probability profiles, calculated as outlined in chapter 1 and reference 100, are depicted in figure 13. There is extensive overlap of the profiles computed for glucagon, secretin, and vasoactive intestinal peptide when the parameters used are those appropriate for water as the solvent. The profiles increase, and become distinguishable from one another, when the parameters appropriate for dodecyl sulfate are substituted. The region of highest helix propagation probability in dodecyl sulfate starts at residue 13 for all three peptides. The helix propagation probability remains high until residue 20, and undergoes a dramatic decrease between residues 20 and 21 for secretin and vasoactive intestinal peptide. In the case of glucagon, however, the helix propagation probability profile undergoes a rather gradual decrease as one proceeds from amino acid

Figure 13. Computed helix propagation probability profiles for glucagon, secretin, and vasoactive intestinal peptide in water (dashed lines), and in the presence of dodecyl sulfate (solid lines). Curves are denoted as G (glucagon), S (secretin), and V (vasoactive intestinal peptide). ¹⁴³



residue 16 towards the carboxyl terminus.

Wu and Yang studied intact glucagon and three tryptic fragments in the presence of dodecyl sulfate.¹⁴⁶ Ordered structure was induced in neither fragment 1-12 nor 13-17. Figure 13 shows that the first 12 residues of glucagon are predicted to experience little effect. While amino acid residues 13-17 in intact glucagon are predicted to become more helical in dodecyl sulfate, the isolated pentapeptide may be too small to experience a conformational change in the presence of sodium dodecyl sulfate. Fragment 19-29 was found to experience a conformational change in the presence of dodecyl sulfate. While this portion of the glucagon molecule is predicted to become more helical in the presence of dodecyl sulfate, the enhanced helicity of the isolated fragment is somewhat puzzling unless aggregation occurs.

No biological activity is exhibited by fragments 1-21, 22-29, 20-29, or 2-29, although the longest of these fragments does bind to the receptor.¹⁴⁷ The three shorter fragments, which show no binding activity, do not retain the region of high helix propagation probability depicted in figure 13 intact. In the crystalline state, amino acid residues 10-25 are found to be helical.¹⁴⁵ Blundell and his coworkers suggest that this segment may also be helical when glucagon interacts with its receptor. The empirical method of Chou and Fasman has also been used to predict several structures for glucagon, one of which has a helical segment at amino acid residues 19-27.¹⁴⁸ The value of $[\theta]_{222}$ for intact glucagon in sodium dodecyl sulfate becomes more negative when the pH is lowered to about 2. In glucagon there are three aspartic acid residues interspersed among the cationic residues which are predicted to interact most strongly with the anionic

detergent. When ionized, these aspartyl residues should produce an unfavorable interaction with the head group of an anionic lipid or detergent. This interaction may explain the pH dependence and also contribute to the thermal lability of the helical structure induced in glucagon by phosphatidic acid and phosphatidyl glycerol.

Secretin

The main differences in the helix propagation probability profiles for glucagon and secretin in dodecyl sulfate lie in the distinctness of the area of maximum probability and in the height of the maximum. Clearly the induced helical segment in secretin is localized in the area of amino acid residues 13-20. The region of high helix propagation probability in figure 13 is in reasonable accord with results of a study of the propensity for chloroethanol-induced helix formation in secretin fragments.¹³⁵ Fragments 15-27, 14-27, and 5-27 showed a propensity for helix formation, which was comparable to that of the intact hormone. On the other hand, fragments 1-14, 18-27, 19-27, 20-27, and 22-27 had a much reduced tendency for helix formation. Those peptides in which helicity is most easily induced retain intact, or nearly intact, the region of high helix propagation probability in figure 13. Another analysis,¹³⁶ using empirical parameters, came to the conclusion that the most probable site for helix formation was at residues 17-24.

The amino terminal portion of secretin, especially residues 1-6, is required for adenylate cyclase activation.¹⁴⁹ Fragment 1-14 can increase cyclic adenosine monophosphate production and has a measurable binding affinity, while fragment 5-27 can bind to secretin

receptors, but exhibits no agonist characteristics. ^{123,124,150} Both fragments 14-27 and 15-27 are competitive inhibitors of secretin binding in pancreatic acinar cells. ^{124,150} From these observations it has been suggested that the amino terminal portion contains the adenylate cyclase activating sequence, while the highest binding affinity lies somewhere towards the carboxyl portion of the chain. The lipid induced helical segment depicted in figure 13 could therefore play a role in peptide-receptor interaction.

Vasoactive Intestinal Peptide

The helix propagation probability profile for vasoactive intestinal peptide is quite dramatic. Amino acid residues 13-20 have an extremely high helix propagation probability. A major difference between secretin and vasoactive intestinal peptide is the absence of anionic residues near the area of highest helix propagation probability in the latter. This feature appears to be one of the main distinguishing factors between the activities of the two hormones. ¹²³ Neither fragment 1-10 nor 18-28 has any effect on the binding of vasoactive intestinal peptide to rat membranes, even at concentrations which are high relative to the concentration of vasoactive intestinal peptide.¹²⁸ Fragments 14-28 and 15-28, however, have a low but recognizable affinity for the high affinity vasoactive intestinal peptide binding sites in pancreatic acinar cells. ^{123,150} Thus the lipid induced helical segment predicted in figure 13 for vasoactive intestinal peptide occurs in an area of the chain which is implicated in the binding process.

CHAPTER 4

β -SHEET TO α -HELIX TRANSITION IN THE BINDING SUBUNIT OF CHOLERA TOXIN

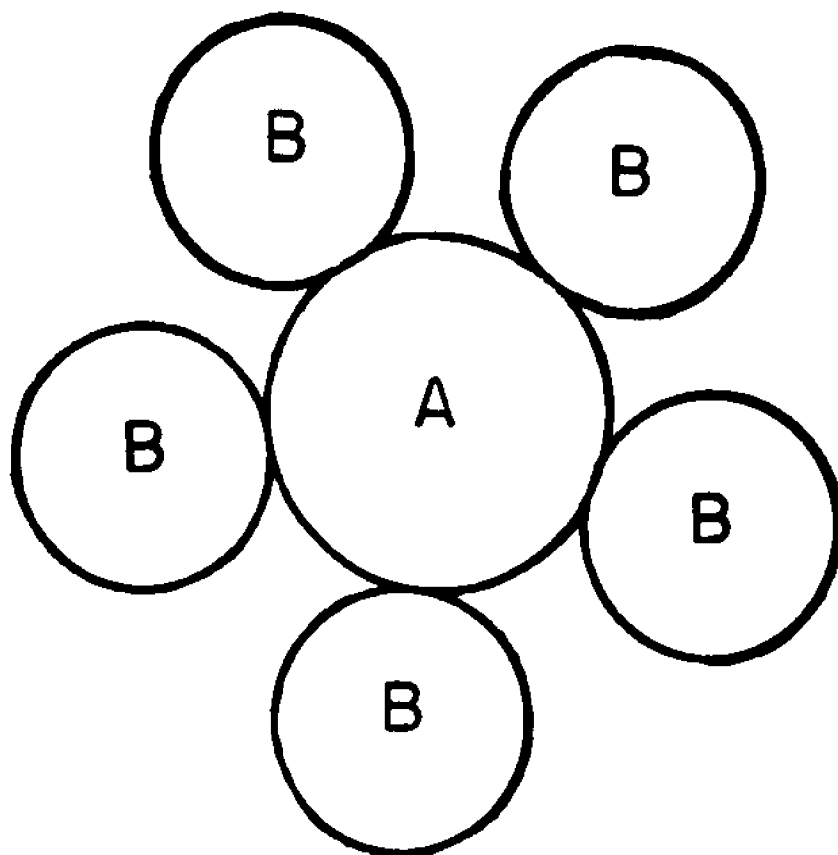
INTRODUCTION

The bacteria Vibrio cholera was isolated by Robert Koch in 1883, and identified as the pathogen of cholera. He suggested that the diarrhea caused by the organism was due to an excreted poison. This was based on the fact that the organism was found only in the intestine and not in any other surrounding tissues. This contention was not borne out by experimentation for many years. The reason for this was the great difficulty encountered in trying to find a suitable animal model for cholera. The first such successful model was developed in 1953.¹⁵¹ A lot of experimentation has been done in recent years with canines,¹⁵² and work is now done with isolated tissues and in cell culture as well.¹⁵³⁻¹⁵⁸

The isolation of pure cholera toxin was first accomplished by Finkelstein and LoSpalluto in 1969.¹⁵⁹ This led to great advances in present day understanding of the structure and mode of action of the toxin. The toxin is composed of two different types of subunits, A and B.^{160,161} The A subunit is composed of two chains which are linked by a disulfide bond. The A subunit is the substituent of the toxin which elicits the characteristic changes in the target cells, including the activation of adenylate cyclase.¹⁶² The binding or B subunit is responsible for the binding of the toxin to the cell membrane via specific interaction with ganglioside G_{M1} .¹⁶³⁻¹⁶⁷ The toxin has the subunit structure AB_5 as depicted in figure 14.¹⁶¹

The purpose of this chapter is to describe a novel conformational change induced in subunit B by sodium dodecyl sulfate, which is

Figure 14. A depiction of the quaternary structure of the cholera toxin. The structure AB_5 is given as proposed to account for the equivalence of the binding subunits.



possibly related to the pathogenesis of the cholera toxin. The B subunit has a molecular weight of 11,604.¹⁶⁸ The amino acid sequence is shown in figure 15.¹⁶⁸ Arginyl 35 has been implicated in the binding process of the B subunit with ganglioside G_{M1} .¹⁶⁹ In so far as the lipid is concerned, the specificity of the interaction apparently resides in the oligosaccharide portion of the molecule.¹⁷⁰ There is a lag time of several hours between binding of the toxin and fluid loss.¹⁵² It has been proposed that perhaps this lag arises from the necessity for a conformational change in the toxin.¹⁶⁵

Experimental circular dichroism studies and the previously described statistical mechanical methods have been combined successfully to rationalize the conformational changes produced in endogenous opioid peptides and hormonal proteins upon their interaction with anionic lipids.^{129,130,143,171} When these methods are applied to the B subunit of cholera toxin, these methods predict a conformation in the presence of anionic lipid which is quite different from the structure thought to exist in water.¹⁷⁰ If the ganglioside G_{M1} -cell membrane environment bears a similarity to that provided by the anionic lipid, then the conformation of subunit B may indeed change upon binding. Circular dichroism studies described here provide experimental confirmation for the major conclusions from theory. The conformational change observed was found to take place over a period of time comparable to the lag period observed between toxin binding and fluid loss. The increase in helicity along with other theoretical and structural factors point towards a helical hairpin structure for the B subunit in dodecyl sulfate.

Figure 15. The amino acid sequence for the binding subunit of cholera toxin.¹⁶⁸ The two cysteine residues involved in disulfide bond formation are underlined as well as the sequences Pro-Gly-Ser (PGS), which has a high probability for reverse turn formation.

10	20	30	40	50
TFQNI DL CA	EYHNTQIHTL	NNKIFSYTES	LAGKREMAII	TFKNGATFEV
60	70	80	90	100
EVPGSQHIDS	QKKAIERMKN	TLRIAYLTEA	KVEKLCVWNN	KTPHAIAAIS

103

MAN

MATERIALS AND METHODS

Subunit B of cholera toxin from Vibrio cholera and dithiothreitol were obtained from the Sigma Chemical Company. Sodium dodecyl sulfate was obtained from Matheson, Coleman, and Bell Manufacturing Chemists.

Protein solutions used in this investigation had an approximate concentration of 0.0187 mg/ml. The concentration of the subunit B was determined from the absorbance and extinction coefficient at 280 nanometers.¹⁷² Circular dichroism measurements were made using a Durrum-Jasco recording spectropolarimeter calibrated with d-10-camphor-sulfonic acid.¹³¹ Fused quartz cells with a pathlength of 1 and 10 mm were used. Cell temperature was maintained with a brass water jacket surrounding the cell which was connected to a constant temperature water bath. Sample temperature was measured with a Yellow Springs Instruments model Y2Sc telethermometer before and after each scan.

The theoretical and experimental basis for the statistical mechanical calculations was described in chapter one and elsewhere.¹⁰¹⁻²

RESULTS

The circular dichroism of the B subunit in water is negative throughout the spectral range studied (figure 16). A single minimum was found near 215 nm having a mean residue ellipticity of $-5000 \text{ deg cm}^2 \text{ dmol}^{-1}$. The shape and intensity are in reasonable agreement with spectra reported earlier for the B subunit.^{169,170} The spectrum in water is indicative of a virtual absence of alpha helix and the presence of appreciable beta-structure.

In the presence of dodecyl sulfate, the circular dichroism changes drastically over a period of several hours. There is an increase in intensity and a gradual development of the double minima which are characteristic of the alpha helix. The equilibrium spectrum is shown in figure 16. The time dependence of the mean residue molar ellipticity at 222 nm is depicted in figure 17. Half of the change is completed in about twenty minutes, but roughly ten times as long is required to attain equilibrium. The profile suggests that multiple steps are involved in the rearrangement. The equilibrium spectrum is less intense by about 15 % in the presence of excess dithiothreitol. The development of helicity is then favored by the presence of an intact disulfide bond between cysteine residues 9 and 86. The temperature dependence of the equilibrium spectrum is smaller when the disulfide bond is intact. For the case of the disulfide bond broken and intact the values of $d[\theta]_{222}/dT$ are 84 and 54 $\text{deg cm}^2 \text{ dmol}^{-1} \text{ K}^{-1}$, respectively. The value of $[\theta]_{222}$ obtained in dodecyl sulfate suggests a helical content of approximately 33 %. A helical content near zero is deduced from the

Figure 16. Circular dichroism of subunit B in (A) water, and (B) 0.003 M dodecyl sulfate. Conditions were pH 7.0 at 30°C and pH 7.04 at 25°C, respectively.

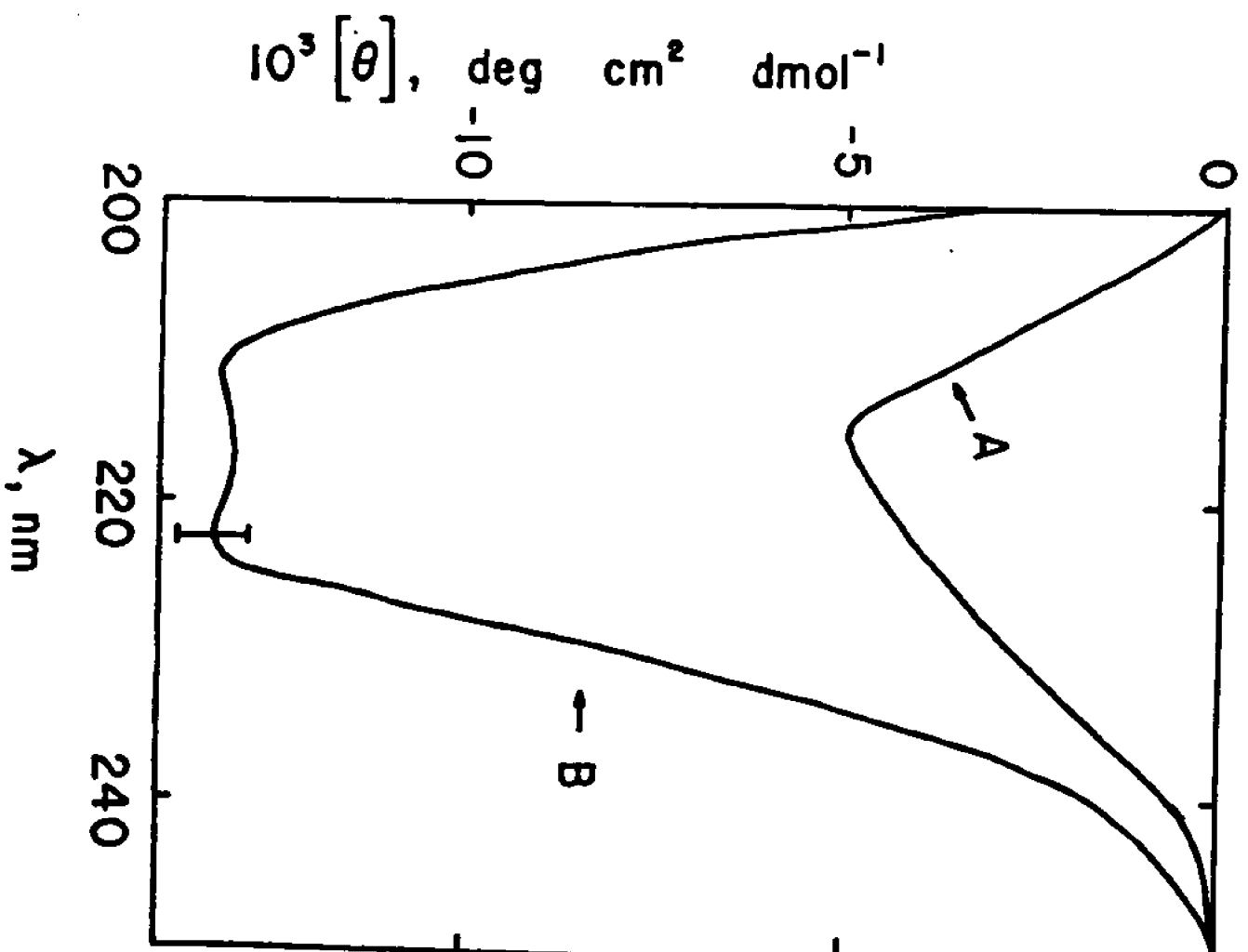
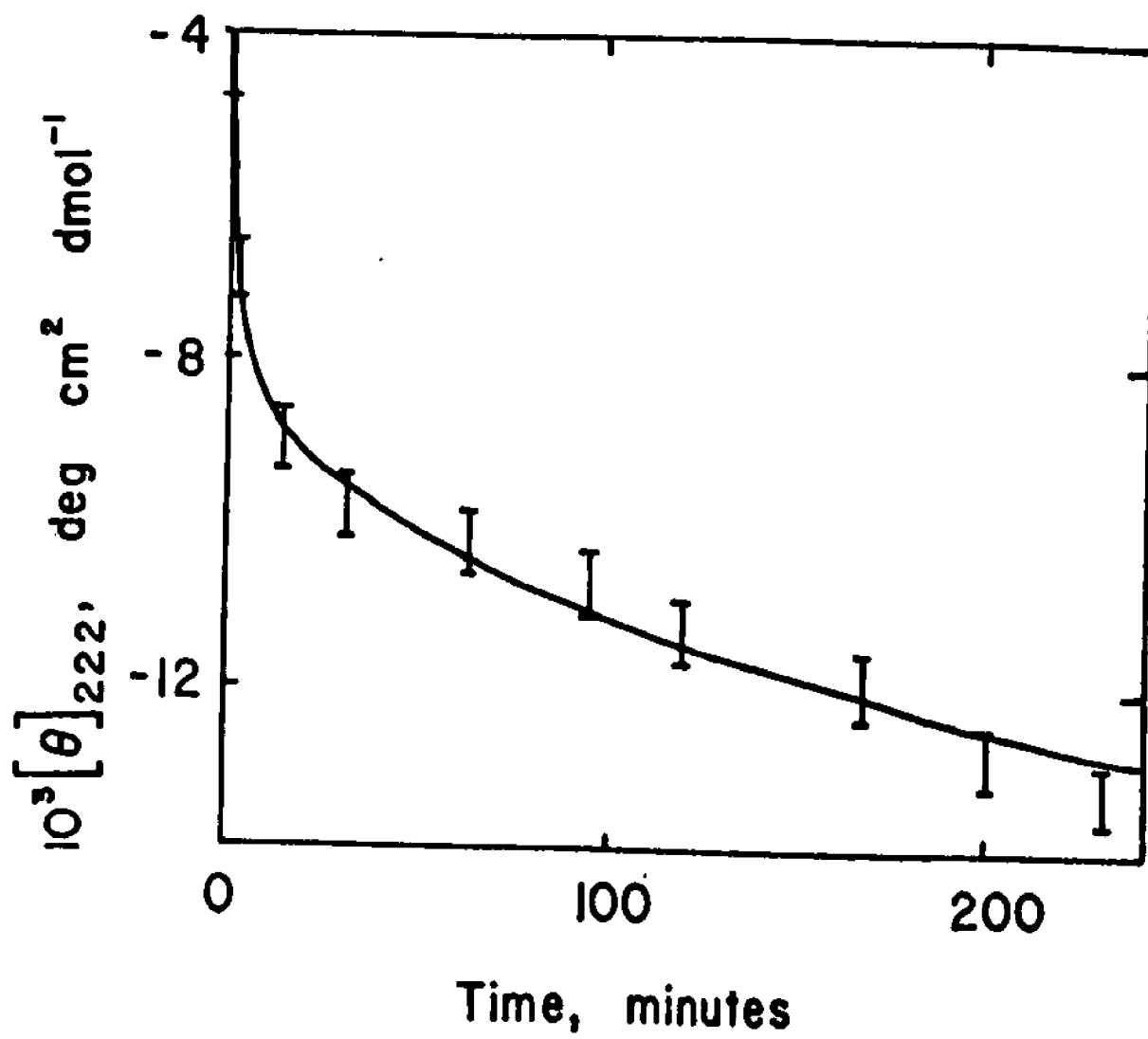


Figure 17. Time dependence of $[\theta]_{222}$ for subunit B in 0.003 M dodecyl sulfate, pH 7.04, 25°C. The line through the points was calculated for two consecutive first-order reactions. Rate constants were $k_1 = 0.21 \text{ min}^{-1}$ and $k_2 = 0.0069 \text{ min}^{-1}$ and the ellipticities of the three species, in order of their appearance, were -4,400, -8,700, and -13,900 $\text{deg cm}^2 \text{ dmol}^{-1}$.



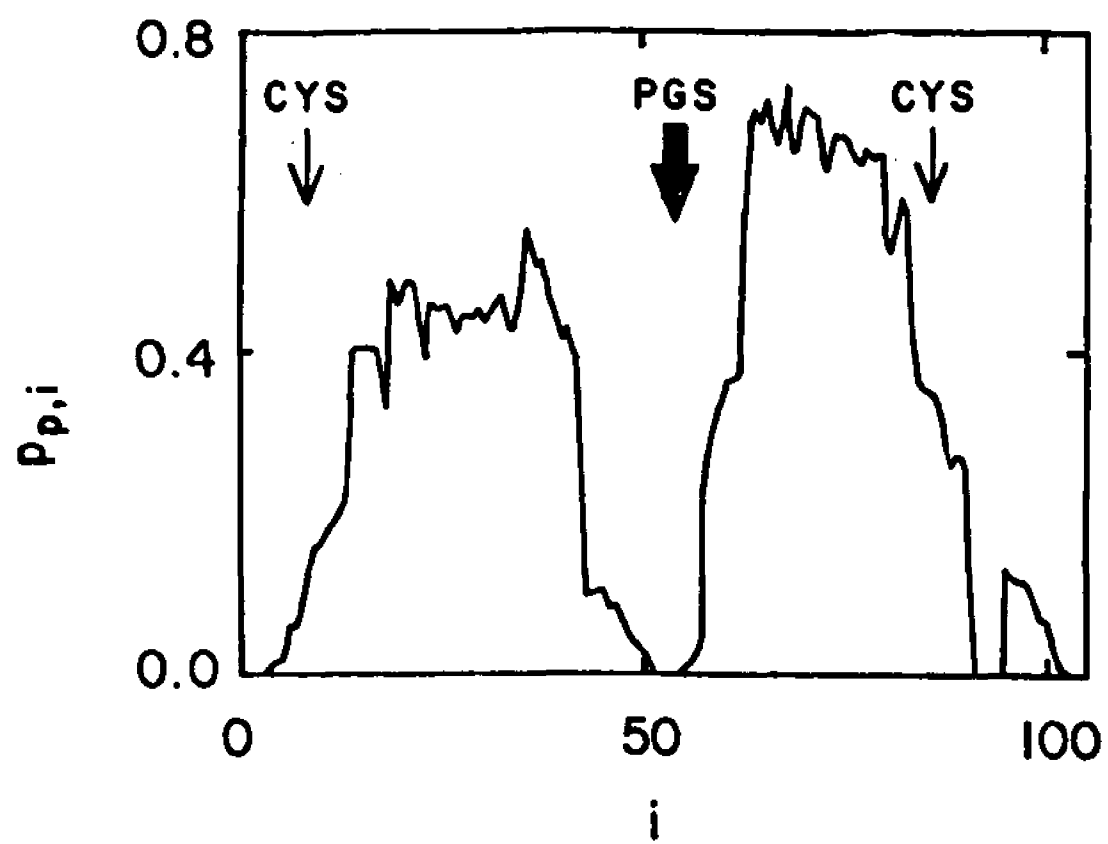
circular dichroism in water. The matrix calculation predicts helical contents of 8 % in water and 38 % in the presence of dodecyl sulfate. ^{143,100} Experimental circular dichroism spectra change in manner predicted by the theory.

DISCUSSION

The theoretical prediction (8 % helix in the absence of dodecyl sulfate, and 38 % in the presence of dodecyl sulfate) is in reasonable agreement with the circular dichroism shown in figure 16. Agreement between observed and calculated helicities for the native state may be, in part, fortuitous. The calculation does not take into account long range interactions, which are undoubtedly important in the determination of the conformation of the native structure for this protein. Greater significance should be attached to the agreement between theory and experiment in so far as the helicity in the presence of dodecyl sulfate. Since the configuration partition function successfully reproduces the fraction helix under these conditions, it is reasonable to extract information from this partition function, such as the helix propagation probability profile (figure 18) for this molecule in the presence of dodecyl sulfate. The formation of two helical segments containing as many as thirty residues each is clearly predicted by this calculation. Experimental circular dichroism shows slightly more helical structure (33 % versus 28 %) is induced when the disulfide bond is intact than in the presence of excess dithiothreitol.

Residues 53-55, Pro-Gly-Ser, have a high probability for reverse turn formation.^{89,173} This tripeptide is found in a reverse turn in adenylate kinase (residues 17-19).⁸⁹ A simple model of the B subunit in dodecyl sulfate might be a helical hairpin, with the PGS sequence forming a reverse turn at the head of the hairpin, and

Figure 18. Computed helix propagation probability profile for subunit B in the presence of dodecyl sulfate. Locations of the cysteinyl residues involved in disulfide bond formation, as well as the Pro-Gly-Ser (PGS) sequence, are noted.¹²⁹

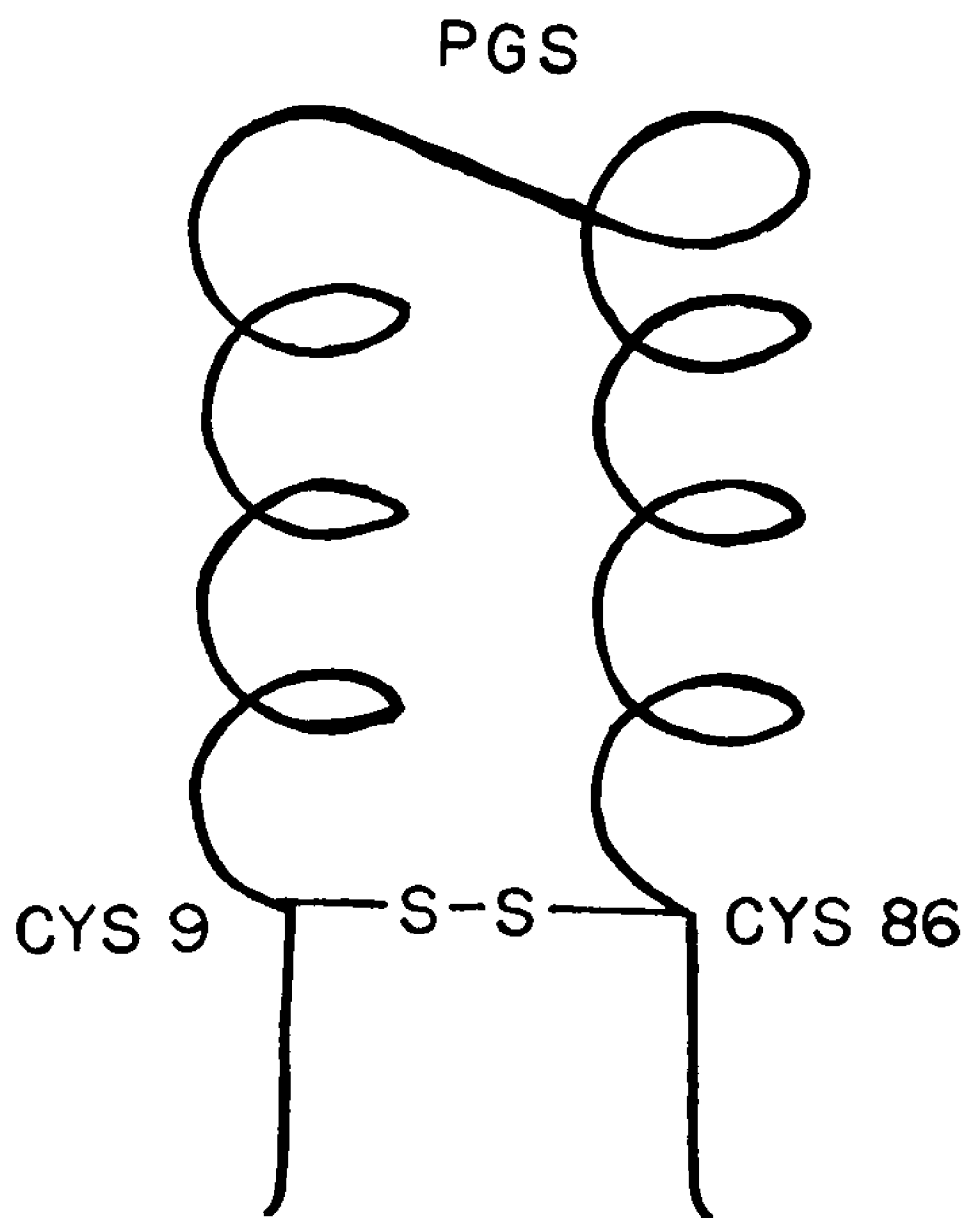


the disulfide bond crosslinking the ends. The proposed conformation is depicted in figure 19. Helix-helix interaction^{174,175} could occur in such a structure along much of the length of the hairpin. Rupture of the disulfide bond would make accessible conformations in which the helices are sufficiently remote from one another that stabilizing helix-helix interactions are negated. The helical content would then decrease upon reduction of the disulfide bond, in harmony with the response of the circular dichroism spectra to excess dithiothreitol.

Such a hairpin conformation may be important for the biological function of the B subunit. The cellular receptor for subunit B is ganglioside G_{M1} ,¹⁶³⁻¹⁶⁷ an anionic amphipathic molecule. In studies of cholera in canines,¹⁵² Carpenter and coworkers observed that the onset of fluid loss through the intestines always follows a lag period of several hours. Sterile filtrates of Vibrio cholera were usually absorbed by the jejunal loops within an hour, and fluid secretion started after two hours, and reached a maximal rate at four hours, which lasted four to six hours. A lag period of 15-20 minutes has been observed in adenylate cyclase activation studies with cholera toxin in isolated tissues and cells.¹⁵³⁻¹⁵⁸ From figure 17 it can be seen that the time scale of the conformational change in dodecyl sulfate is comparable to that of the lag period.

Further study of this conformational transition is warranted for several reasons : (i) It is an excellent model for a β -structure to α -helix transition in a protein. (ii) The molecule constitutes a large macrocycle, closed by a disulfide bond, which provides for helix-helix interaction. (iii) The conformational change may be related to the means by which subunit B facilitates the penetration of subunit A into

Figure 19. A depiction of the proposed helical hairpin structure for the subunit B in the presence of sodium dodecyl sulfate.¹²⁹



the cell.

PART TWO

STUDIES ON SYNTHETIC RANDOM COPOLYPEPTIDES

CHAPTER 5

HELIX INITIATION AND PROPAGATION BY HYDROXYETHYL-L-GLUTAMINYL RESIDUES IN WATER

INTRODUCTION

The series of homopolypeptides based on N⁵-(ω -hydroxyalkyl)-L-glutamine has been studied extensively. The three most prominent members of this series are the poly(hydroxyalkyl)-L-glutamines with alkyl = ethyl, propyl, and butyl. These three water soluble nonionic homopolypeptides exhibit a range of conformational properties. The hydrodynamic ¹⁷⁶ and optical ¹⁷⁶⁻¹⁷⁸ properties of poly(hydroxyethyl)-L-glutamine in water are those expected for a disordered homopolypeptide bearing a -CH₂R sidechain. In contrast, near 0°C in water poly(hydroxybutyl)-L-glutamine is nearly completely helical and poly(hydroxypropyl)-L-glutamine is partially helical. ^{33,108,133,179-80} The helical content of these two polymers is reduced by an elevation in temperature. The helicity of all three poly(hydroxyalkyl)-L-glutamines can also be modified by changing the solvent composition. Both inorganic salts ^{108,133,181} and organic cosolvents ^{108,177,179,180,182-7} are effective in this regard.

The Zimm-Bragg ²⁸ helix-coil stability constants, σ and s , for the hydroxypropyl-L-glutaminy ^{33,179,180} and hydroxybutyl-L-glutaminy ³³ residues in water have been determined from studies of the homopolypeptides and also via the host guest technique. In this chapter the determination of σ and s for the hydroxyethyl-L-glutaminy residue in water is described. They were determined using the host guest technique, with hydroxybutyl-L-glutaminy residues playing the role of host. When the values determined in this chapter are examined in conjunction with the previously determined σ and s values for

hydroxypropyl-L-glutaminy³³ and hydroxybutyl-L-glutaminy³³ residues, then the results provide insight into the conformational consequences of the addition of methylene groups to the periphery of an amino acid residue sidechain.

EXPERIMENTAL

Copolypeptides were prepared by Dr. Erin R. Hawkins by aminolysis of poly(γ -benzyl-L-glutamate) with a mixture of hydroxyethanolamine and hydroxybutanolamine. Reaction conditions were a slight modification of the procedure by which poly(hydroxybutyl-L-glutamine) was prepared from poly(γ -benzyl-L-glutamate).¹⁸⁸ Compositions were determined from a quantitative ninhydrin analysis of the hydrolysate, using an amino acid analyzer, and also from proton nuclear magnetic resonance spectra. Similar results were obtained by both methods. The hydroxyethanolamine/hydroxybutanolamine ratio in the hydrolysate was found to be significantly higher than that in the reaction mixture used to prepare the copolymers.

Circular dichroism measurements were conducted on solutions prepared by dilution of a stock solution with distilled deionized water. Circular dichroism measurements were made with a Durrum-Jasco J-20 recording spectropolarimeter calibrated with d-10 camphorsulfonic acid in water.¹³¹ Measurements were made using fused quartz cells with a pathlength of 10 mm. Cell temperature was maintained with a circulating water bath connected to a brass water jacket surrounding the cell. Sample temperature was determined with a Yellow Springs Instruments model Y2Sc telethermometer before and after each spectrum was recorded. Solution concentrations were determined from a micro-Kjeldahl analysis of the stock solutions.¹⁸⁹

CALCULATIONS

Calculations were carried out using a sample of random copolypeptide chains generated with the aid of a pseudo random number generator. It is assumed that the conformational properties of the chains are dominated by nearest neighbor interactions, the polymerization process results in truly random copolypeptides, and end effects are treated sufficiently using the 2×2 statistical weight matrix.

The statistical weight matrix for amino acid residue i is ;

$$\underline{U}_i = \begin{bmatrix} 1 & \sigma s \\ 1 & s \end{bmatrix}_i \quad (85)$$

where rows index the state of residue $i-1$, columns index the state of amino acid residue i , and the order of indexing is c, h . An amino acid residue whose conformation is that found in the alpha helix is in state h , and all other conformations are described by state c . The statistical weight for propagation of a helical segment is s , and the statistical weight for initiation of a helical segment is σs . If there is compensation between long range intramolecular interactions and solvent-macromolecule interactions, the configuration partition function, Z , for a partially helical copolypeptide containing n amino acid residues is ;

$$Z = \underline{J}^* \underline{U}_1 \underline{U}_2 \dots \underline{U}_n \underline{J} \quad (86)$$

where \underline{J}^* is $\text{row}(1,0)$ and \underline{J} is $\text{col}(1,1)$. The fraction helix, f_h , is obtained as ;

$$f_h = n^{-1} Z^{-1} \underline{J}^* \hat{\underline{U}}_1 \hat{\underline{U}}_2 \dots \hat{\underline{U}}_n \underline{J} \quad (87)$$

where \underline{J}^* is now row(1,0,0,0) and \underline{J} is col(0,0,1,1). The supermatrix $\hat{\underline{U}}_i$ is ;

$$\hat{\underline{U}}_i = \begin{bmatrix} \underline{U} & \underline{U}' \\ 0 & \underline{U} \end{bmatrix}_i \quad (88)$$

where zero denotes a null matrix and \underline{U}' is obtained by zeroing all elements in the first column of \underline{U}_i .

Throughout the 1960's several workers developed methods which could be used to approximate the helix-coil stability constants in water from the pertinent melting curves. Two highly visible examples were those developed by Lifson¹⁹⁰ and Allegra.¹⁹¹ These two models were combined and modified by Poland and Scheraga to form the Lifson-Allegra-Poland-Scheraga or "LAPS" hierarchy of approximations.³⁴ This is the calculational technique which has been used quite successfully by Scheraga and his coworkers as described in the introduction.³⁵⁻⁵¹ These methods were developed as approximations which allowed the determination of σ and s for a given residue within the severe limitations of the computational capabilities of the computers in use at the time. The explicit calculation as outlined above is no longer exorbitant in required computer time and was therefore used as described below.

This calculation was carried out explicitly for the random copolypeptides as follows. A sequence of random numbers, each between zero and one, was generated. If the number was less than or equal to the fraction of hydroxyethyl-L-glutaminy residues in the copolypep-

tide, the residue in question was designated as hydroxyethyl-L-glutaminyl. Otherwise the residue was hydroxybutyl-L-glutaminyl. A value for n of 300 was used as this gave a good approximation to a chain of infinite molecular weight without using exorbitant amounts of central processing unit (CPU) time. The helical content of this chain was then calculated from equation 87. The desired helical content is the average of such values over an infinitely large sample of random copolypeptides. For present purposes this helical content was approximated as the average of values obtained from five sets, each containing 100 independently grown chains. This sample size was deemed sufficiently large because there was good agreement on the averages deduced from the five sets.

Values of σ and s for hydroxybutyl-L-glutaminyl residues were taken from von Dreele et al.³³ Those for hydroxyethyl-L-glutaminyl residues were obtained from the best fit to the experimentally determined values of f_h for the copolypeptides studied.

RESULTS

Figure 20 depicts f_h as a function of hydroxybutyl-L-glutaminyl composition at several temperatures. The points are from experiment, and the lines are the fit obtained from the theory. Copolypeptides whose compositions are 37 % and 39 % hydroxybutyl-L-glutaminyl give discordant results, particularly at lower temperatures. The anomaly appears to reside in the copolymer which is 37 % hydroxybutyl-L-glutaminyl because the 39 % hydroxybutyl-L-glutaminyl copolypeptide gives results which fall on the trend established by the other copolypeptides and the two homopolypeptides. For this reason the 37 % hydroxybutyl-L-glutaminyl copolypeptide was ignored in the analysis.

Figure 21 depicts the temperature dependence of the values of s deduced for the hydroxyethyl-L-glutaminyl residue when σ for this residue is taken to be 1×10^{-5} . The relative error in s for the hydroxyethyl-L-glutaminyl residue is largest at the highest temperature because the copolypeptides have low helical contents under these conditions. The relative error at 25°C is indicated in figure 21.

Values of s for the hydroxypropyl-L-glutaminyl and hydroxybutyl-L-glutaminyl residues, obtained from von Dreele et al, ³³ are also shown for comparison in figure 21. The values of s for these residues are larger than those for hydroxyethyl-L-glutaminyl, as expected. In all three cases s decreases as temperature increases. While the temperature effect is smallest with hydroxyethyl-L-glutaminyl, the form of the thermally induced change is similar for the three hydroxyalkyl-L-glutaminyl residues. The s for the L-glutaminyl residue, from Denton et al., ¹⁹² also decreases with an increase in temperature, but the

Figure 20. Experimentally determined f_h as a function of hydroxybutyl-L-glutaminyl composition at temperatures of 4 (●), 25 (▲), 34 (▼), 45 (■), 56 (◆), and 64 (†) °C. Solid lines are calculated from theory.

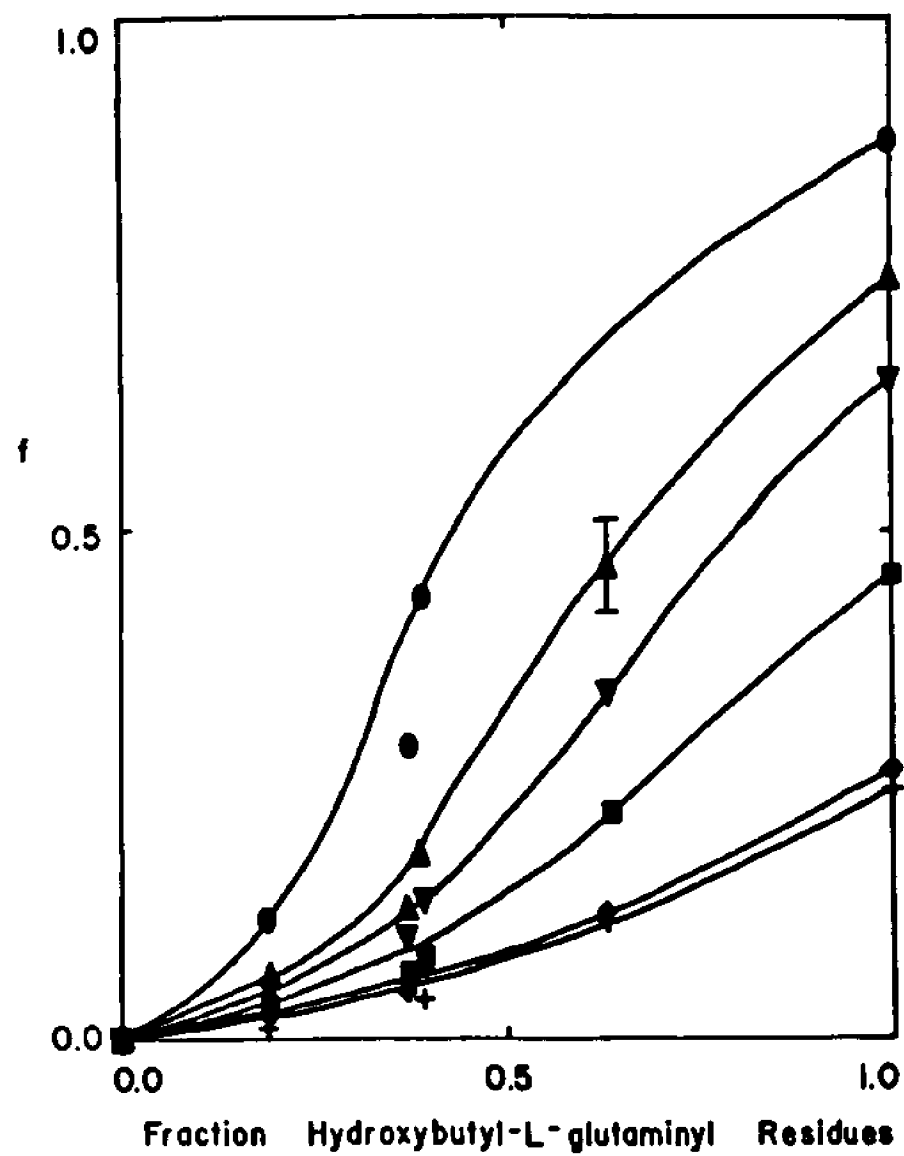
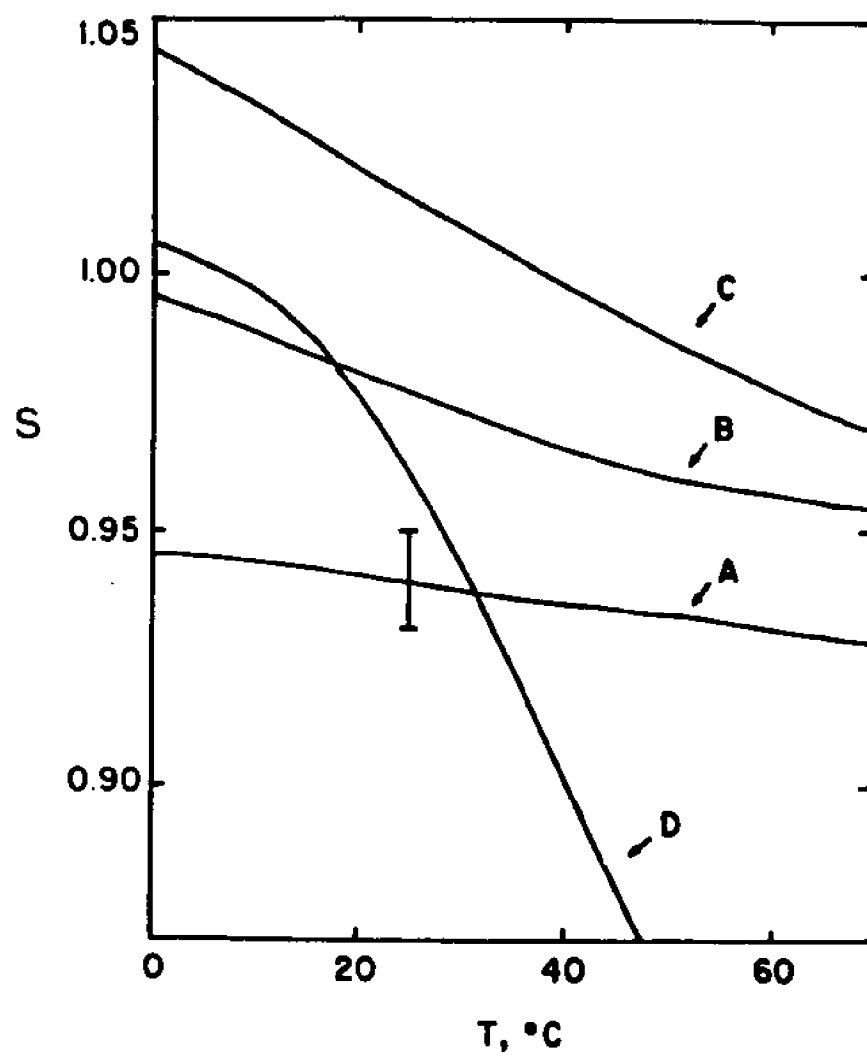


Figure 21. The parameter s as a function of temperature for hydroxyethyl-L-glutaminyl (A), hydroxypropyl-L-glutaminyl (B), hydroxybutyl-L-glutaminyl (C), and L-glutaminyl residues. Curves B and C are from von Dreele et al.,³³ and curve D is from Denton et al.¹⁹²



form of the temperature dependence is rather different from that seen with the hydroxyalkyl-L-glutaminyI residues. Specifically, d^2s/dT^2 is strongly negative at the lower end of the temperature range for L-glutaminyI, while it is zero or even slightly positive for the hydroxyalkyl-L-glutaminyI residues.

DISCUSSION

The ability to theoretically describe the experimentally determined f_h for the copolypeptides is only weakly dependent upon the value assigned to σ for the hydroxyethyl-L-glutaminyl residue. Results presented here were obtained by using $\sigma = 1 \times 10^{-5}$. Insensitivity of the calculation to σ for the hydroxyethyl-L-glutaminyl residue can be attributed to the fact that it is much smaller than the value of σ for the host residue, hydroxybutyl-L-glutaminyl (table 3). Nearly all helix initiation occurs at hydroxybutyl-L-glutaminyl residues, causing the value assigned to σ for the hydroxyethyl-L-glutaminyl residue to be of little consequence so long as it is small.

The value of σ for the hydroxyethyl-L-glutaminyl residue is much smaller than the value of $(1.5-4.4) \times 10^{-3}$ estimated in water-2-propanol mixtures.¹⁹³ It is also smaller than the σ found by von Dreele et al.³³ for hydroxypropyl-L-glutaminyl and hydroxybutyl-L-glutaminyl residues in water. (table 3). However, it follows qualitatively the trend established by the other two residues of the series. In this series helix formation becomes less cooperative as methylene groups are added to the periphery of the sidechain. The L-glutaminyl residue has a much larger σ ,¹⁹² and consequently does not behave as a member of the hydroxyalkyl-L-glutaminyl series.

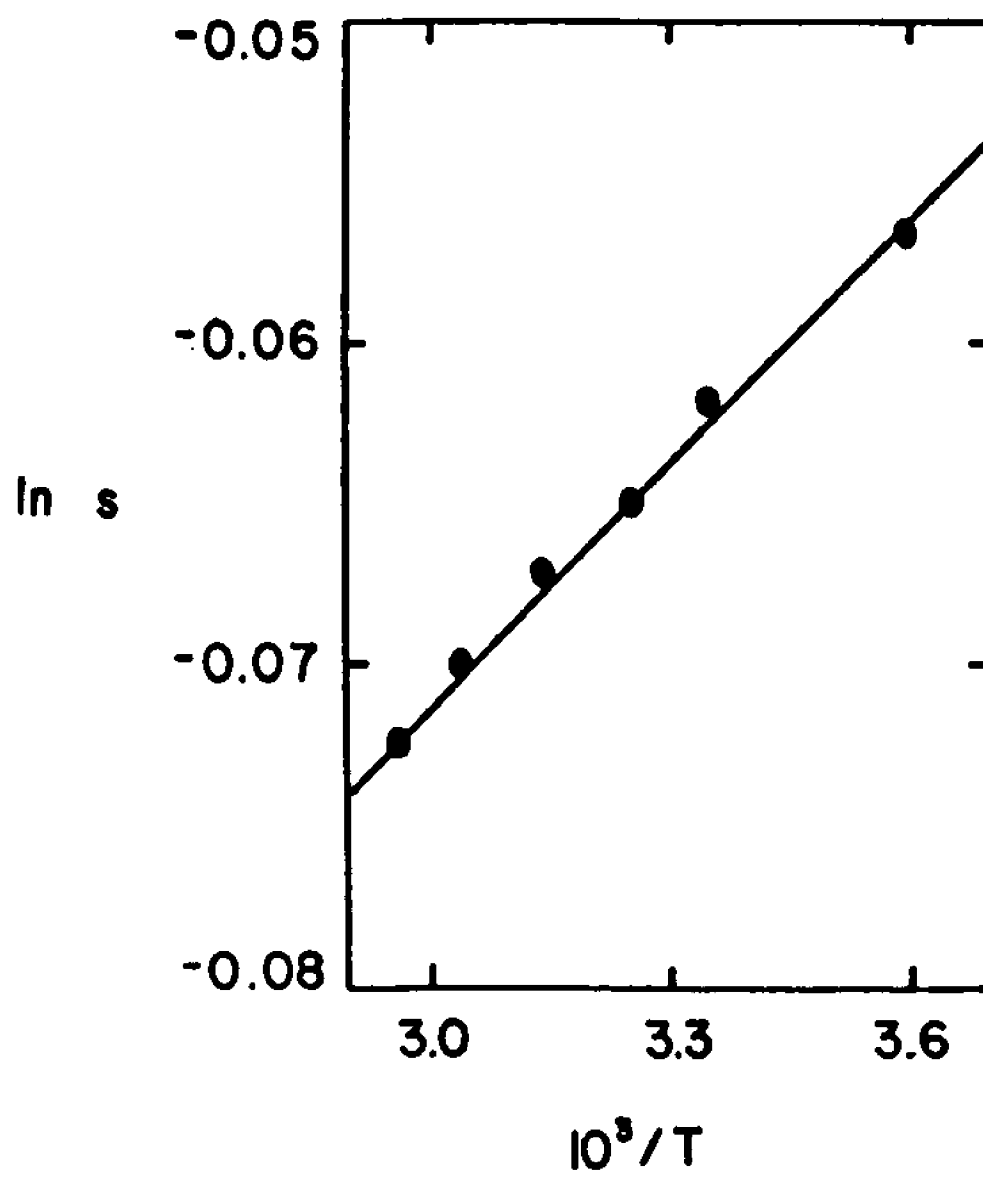
Figure 22 demonstrates that $\ln s$ is a linear function of $1/T$, signifying that ΔH for helix propagation is independent of temperature over the range covered. Thermodynamic parameters for helix propagation at 20°C are collected in table 3 along with the values obtained by Scheraga and coworkers.^{33,192} for the hydroxypropyl-L-glutaminyl,

TABLE 3

Parameters for Random Coil to Helix Transition in Water at 20°C

Parameter	L-glutaminyl ¹⁹²	Hydroxyethyl- L-glutaminyl	Hydroxypropyl- L-glutaminyl ³³	Hydroxybutyl- L-glutaminyl ³³
σ	3.3×10^{-3}	1×10^{-5}	2.2×10^{-4}	6.7×10^{-4}
$\Delta H, \text{ cal mol}^{-1}$	-493	-127	-168	-195
$\Delta G, \text{ cal mol}^{-1}$	13	35	13	-11
$\Delta S, \text{ eu mol}^{-1}$	-1.7	-0.55	-0.616	-0.627

Figure 22. A plot of $\ln s$ versus $1/T$ (correlation coefficient 0.996).



hydroxybutyl-L-glutaminy], and L-glutaminy] residues. In the hydroxy-alkyl-L-glutaminy] series, ΔH becomes more negative by about 35 cal mol^{-1} with the addition of a methylene group, while ΔG at 20°C becomes more negative by about 25 cal mol^{-1} . The thermodynamics of helix propagation for the L-glutaminy] residue is very different from that expected for a hypothetical hydroxyalkyl-L-glutaminy] residue in which the "alkyl" group contains no methylene units.

Conformational energy calculations have been performed for the hydroxyethyl-L-glutaminy] residue in order to determine the preferred conformations.¹⁹³ The sidechain was found to participate in intramolecular hydrogen bond formation between the ω -hydroxyethyl group and the carbonyl group in the peptide backbone. This hydrogen bond can be formed without appreciable alteration in the average dimensions of the chain. When the characteristic ratio is calculated including the consequences of this interaction,¹⁹³ it remains in agreement with the experimental result.¹⁷⁶ Thus the average dimensions do not provide a sensitive test for the importance of intramolecular hydrogen bond formation by this residue in water.

Formation of an intramolecular hydrogen bond involving the carbonyl group in the backbone would discourage helix formation. If this intramolecular hydrogen bond were relatively inaccessible to higher hydroxyalkyl-L-glutaminy] residues, helix formation by hydroxyethyl-L-glutaminy] might be expected to be more difficult than that predicted from the observed behavior of hydroxypropyl-L-glutaminy] and hydroxybutyl-L-glutaminy] residues. Results reported in figure 21 and table 3 show that there are no anomalies in the behavior of s for the hydroxyethyl-L-glutaminy] residue. The intramolecular hydrogen bonded conformations

predicted for the hydroxyethyl-L-glutaminyl may be of little consequence when they must compete with a hydrogen bonding solvent such as water.

CHAPTER 6

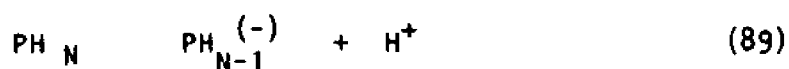
ELECTROSTATIC INTERACTIONS IN RANDOM COPOLYPEPTIDES CONTAINING L-GLUTAMYL RESIDUES

INTRODUCTION

One needs only to examine the structure of any protein or nucleic acid molecule to see at once the importance of electrostatic interactions in so far as the conformational properties of these molecules are concerned. This chapter deals with the treatment of electrostatic interactions in polypeptides. The purpose of this chapter is to try to extend the treatment of electrostatic interactions from the anionic homopolypeptide poly-L-glutamic acid to random copolypeptides containing either L-alanyl or L-tyrosyl residues.

Several groups have presented both experimental and theoretical studies concerned with the pH induced helix-coil transition in poly-L-glutamic acid. Among these are Wada ¹⁹⁵, Nagasawa and Holtzer ¹⁹⁶, Olander and Holtzer, ¹⁹⁷ Warashina and Ikegama, ¹⁹⁸ and Santiago et al. ¹⁹⁹ The latter two used the Zimm-Rice theory ²⁰⁰ of the helix-coil transition in charged homopolypeptides to rationalize the conformational properties and potentiometric titration data for poly-L-glutamic acid. Before discussing this treatment, a review of the basics of polyelectrolytes is in order. ²⁰¹

Several factors determine the pH of a polyelectrolyte solution. The ionization of one particular carboxyl group on a totally uncharged polymer chain is represented by



The mixed acidity constant for the above reaction is

$$K_O = \frac{a_H (\text{concentration of polymer with chosen carboxyl ionized})}{(\text{concentration of } PH_N)} \quad (90)$$

where PH_N denotes the completely unionized polyelectrolyte molecule, while $PH_{N-1}^{(-)}$ represents a polyelectrolyte molecule with the chosen carboxyl group ionized.

There is a fundamental difference between this first ionization and those which follow. In order to remove subsequent protons not only must the carboxyl group of interest be ionized, but the proton has to be removed from the electrostatic potential of the other charged groups on the polyelectrolyte molecule. This means that each succeeding ionization becomes progressively more difficult than the preceding one. There is also a statistical contribution to the entropy change of each ionization, including the first one. For a polymer at a given state of charge, there is a change in the number of ways in which the charges can be arranged which accompanies each ionization. This effect must be included although this is the only distinction which can be made between molecules of the same degree of ionization. When these facts are taken into consideration, the pH of a polyelectrolyte solution can be expressed as

$$pH = pK_O - \log \{(1-\alpha)/\alpha\} + (1/2.303 RT) (\partial G_{ion}/\partial \alpha) \quad (91)$$

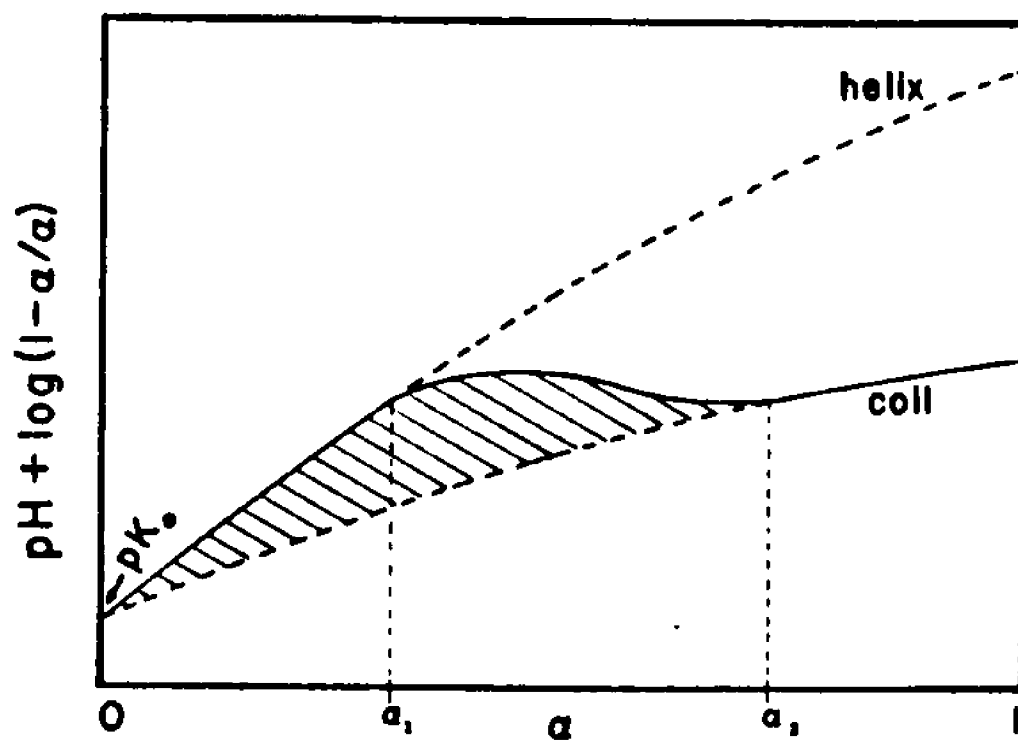
where R is the gas constant, and T is the absolute temperature.

The first term on the right hand side of equation 91 is an intrinsic contribution to every ionization. The second term represents the entropy of each ionization which is related to the increase in the number of ways the charges can be distributed upon each ionization.

In the third term, ΔG_{ion} is the change in electrostatic and conformational free energy which accompanies a change, $\Delta\alpha$, in the degree of ionization. The last term includes both the electrostatic work of removing a proton from the potential of the polyion, and the free energy change associated with any conformational change which accompanies the ionization. If one is dealing with a nonrigid polyelectrolyte, that is one which undergoes a conformational change upon ionization, then the free energy change associated with that conformational change is included in the last term.

From these considerations it is easy to see how the potentiometric titration of such nonrigid polyelectrolytes considered in conjunction with the conformational properties, from circular dichroism as a function of pH, can yield much valuable information. Consideration will now be given to how such data can be used to yield such information. The least complex of the two is the pH induced change in the circular dichroism. When the molecules are fully ionized, the circular dichroism of these molecules exhibit the features of a random coil polypeptide, and as the pH is lowered they undergo drastic changes to almost completely alpha helical molecules. This change occurs over a relatively narrow range of pH, which is consistent with the notion that the process is cooperative in nature. The usual manner of displaying the circular dichroism data is simply the fraction of the residues which are in a helical state versus pH. On the other hand, the usual way of plotting the potentiometric titration data is $\text{pH} - \{\log \alpha / (1-\alpha)\}$ versus α . An example of such a plot, which will henceforth be referred to as the titration curve, is depicted in figure 23. When the molecule is unionized it is simply a rodlike alpha helical molecule, but when

Figure 23. A typical titration curve of a polyelectrolyte which undergoes a conformational change between α_1 and α_2 .



$\alpha = \alpha_1$ the completely or nearly completely ordered molecule is noticeably changed to a mixture of ordered and disordered structure, and when $\alpha = \alpha_2$, the molecule no longer contains any detectable ordered structure. Hence two titrations are actually depicted in the figure : the titration of the alpha helical molecule and the titration of the random coil. While the solid line in figure 23 represents an example of an actual titration curve, the dashed lines represent the titration curves which would be obtained if either conformation was predominate throughout the titration. The area between the titration curve and the extrapolated "coil" titration is related to the free energy of the conformational change. One excellent study in this particular area is that of D. Olander for poly-L-glutamic acid.²⁰² Such considerations will only be discussed here in qualitative terms for reasons which will become apparent later in this chapter.

The first attempt at a treatment for this phenomenon using statistical mechanical principles similar to those outlined in the introduction to this dissertation was by B. H. Zimm and S. A. Rice.²⁰⁰ When only nearest neighbors are considered, the statistical weight matrix is a manageable size (4×4). Zimm and Rice presented a method applicable to homopolymers for a "pseudodiagonalization" of the statistical weight matrix for approximating the largest eigenvalue of the matrix, which can then be used to determine any pertinent parameters. This method was used by Warashina and Ikegama,¹⁹⁸ and later by Santiago et al.¹⁹⁹ to rationalize the titration and circular dichroism data for poly-L-glutamic acid. This work serves as the foundation for the work presented here, as they used the pseudodiagonalization procedure as a starting point, and refined their values by performing

the indicated matrix calculations explicitly.

Several other groups have performed experiments similar to those presented here for similar molecules. However, a quantitative rationalization of such data has not been forthcoming. This is attempted here by extending the Zimm-Rice theory to random copolypeptides which contain L-glutamyl residues, and residues which do not ionize in the range of interest; L-alanyl and L-tyrosyl. The data is rationalized within the confines of the Zimm-Rice theory taking into consideration factors which are fundamental to the chemistry and physics of polyelectrolytes.

EXPERIMENTAL

Random copolypeptides with the following residue molar compositions were purchased from Miles-Yeda, Ltd. : L-glutamic acid:L-alanine (65:35) and L-glutamic acid:L-tyrosine (90:10). Sodium chloride and sodium hydroxide were from Baker Chemicals.

Circular dichroism measurements were made on solutions of copolypeptides made by dilution of stock solutions with distilled deionized water and the appropriate amount of sodium chloride to give a final concentration of 0.1 M. Measurements were made with a Durrum-Jasco J-20 recording spectropolarimeter. Fused quartz cells with a pathlength of 10 mm. were used. Measurements were made at 90°, 220°, and 360°C.

Potentiometric titrations were carried out with copolymer solutions with a concentration of 1-3 mg/ml in 0.1 M NaCl. The solution was placed in a jacketed vessel, which was kept at constant temperature by circulating water from a constant temperature water bath, and a stream of nitrogen was passed over the solution to eliminate potential interference from carbon dioxide. Measurements of pH were made with a Beckman 3500 digital pH meter equipped with a combination electrode. Burets with a total discharge volume of 5 ml were used.

Each solution was titrated with 0.1 M HCl prepared from a 1.0 M standard. Solutions were back titrated with 0.1 M NaOH prepared from a saturated solution the same day. Further use of the results were made only when agreement between the two titrations was good. Negligible contributions were found by solvent from blank titrations.

CALCULATIONS

The theoretical analysis of the potentiometric titration and circular dichroism data is achieved using a matrix formulation which is based on the Zimm-Rice theory.²⁰⁰ The treatment here is restricted to nearest neighbors, and shall hence require a 4×4 statistical weight matrix. The matrix is identical to that of Santiago et al.¹⁹⁹

For ionic homopolypeptides of infinite molecular weight, an approximation to the largest eigenvalue of the matrix is readily obtainable. The resulting expression can then be used to obtain values for the four required parameters for the interaction of neighboring charged residues u_{cc} , u_{ch} , u_{hc} , and u_{hh} . This approximate method was the technique used by Warashina and Ikegama.¹⁹⁸ Santiago and coworkers,¹⁹⁹ on the other hand, used these results as crude estimates for the parameters, which they refined by performing the calculations explicitly. In this way they were able to eliminate any error which may have arisen by using the approximate method for diagonalizing the statistical weight matrix. It is these values which are considered as correct here for the L-glutamic acid residue in 0.1 M NaCl at the indicated temperature.

The statistical weight matrix for the ionizable residues is ;

$$U_i = \begin{array}{c} \text{residue } i \\ \text{residue } i-1 \end{array} \begin{array}{c} c(0) \\ c(-) \\ h(0) \\ h(-) \end{array} \begin{array}{cc} c(0) & c(-) \\ \left[\begin{array}{cc} 1 & \gamma \\ 1 & \gamma u_{cc} \\ 1 & \gamma \\ 1 & \gamma u_{hc} \end{array} \right] & \begin{array}{cc} h(0) & h(-) \\ \begin{array}{cc} \sigma s & \gamma \sigma s \\ \sigma s & \gamma \sigma s u_{ch} \\ s & \gamma s \\ s & \gamma s u_{hh} \end{array} \end{array} \end{array} \quad (92)$$

This matrix differs from the one employed by Warashina and Ikegama by the fact that this matrix is the transpose of their matrix. Columns index the state of residue i , and the rows that of $i-1$, and the order of indexing is $c(0)$, $c(-)$, $h(0)$, and $h(-)$. Here h denotes a residue in a helical state, c denotes a residue in a nonhelical or coil state, and the charge state of the sidechain is in parentheses. The Zimm-Bragg²⁸ parameters σ and s here refer to the electrically neutral residue. The intrinsic dissociation constant and pH determine γ according to $\gamma = 10^{\text{pH}-\text{pK}_0}$. The statistical weight matrix appropriate for a residue which is not ionizable in the pH range of interest is simply the matrix in equation 92 with the second and fourth columns nulled. The contribution to the statistical weight made by the electrostatic interaction of neighboring sidechains is denoted by u_{cc} , u_{ch} , u_{hc} , and u_{hh} . The intuitive expectation is that $1 > u_{cc} > u_{hh}$ because electrostatic interactions are repulsive and much more severe in the alpha helix than in the random coil. The relatively infrequent occurrence of helix ends in the charged copolypeptides should permit the use of an approximate form of u_{ch} and u_{hc} . The approximation applied here is $u_{ch} = u_{hc} = (u_{cc} \cdot u_{hh})^{\frac{1}{2}}$.

The configuration partition function, Z , is

$$Z = \text{row}(1,0,0,0) \hat{U}_i^n \text{col}(1,1,1,1) \quad (93)$$

Following the principles described in the introduction, the probability that a residue will be ionized, α , is

$$\alpha = n^{-1} Z^{-1} \text{row}(1,0,0,0,0,0,0,0) \hat{U}_{\alpha}^n \text{col}(0,0,0,0,1,1,1,1) \quad (94)$$

and the probability that a residue is in a helical state is

$$f_h = n^{-1} z^{-1} \text{ row}(1,0,0,0,0,0,0,0) \hat{U}_f^n \text{ col}(0,0,0,0,1,1,1,1) \quad (95)$$

Here, \hat{U}_x is defined as

$$\hat{U}_x = \begin{bmatrix} U_x & U'_x \\ 0 & U_x \end{bmatrix} \quad (96)$$

where U'_x is obtained from U_x by nulling the elements in two columns. The columns nulled are the first and third for $x = \alpha$, and the first and second for $x = f_h$.

The calculation of f_h and α was carried out explicitly for the random copolypeptides as follows : A sequence of random numbers, each between zero and one, was generated. If the number was less than or equal to the mole fraction of L-glutamyl, then the residue in question was designated as L-glutamyl. Otherwise the residue was designated as either L-alanyl or L-tyrosyl. A value of n , the number of residues per chain, of 300 was used as this gives a good approximation to a chain of infinite molecular weight without expending a great deal of computer time. f_h and α were calculated using equations 94 and 95, respectively. The desired α and f_h is the average of such values over an infinitely large sample of random copolypeptide chains. For present purposes these quantities were approximated as the average of such values obtained from four sets, each containing 250 independently grown chains. This sample size was deemed sufficiently large because there was good agreement in the averages deduced from the four sets.

The values of σ and s for the discharged L-glutamyl residue in 0.1 M NaCl are taken from Santiago et al.¹⁹⁹ The values for the L-alanyl and L-tyrosyl residues are from references 36 and 42, respectively.

RESULTS

The circular dichroism spectra for both copolypeptides are those expected for a random coil at neutral pH at all temperatures. Both copolypeptides undergo a random coil to alpha helix transition as the pH is lowered and the carboxyl groups of the L-glutamyl residues are neutralized. The only conformational states consistent with the circular dichroism spectra are the random coil and the alpha helix.

The helical content for the copolypeptides was calculated from the mean residue ellipticity at 222 nanometers as follows ;

$$f_h = ([\theta] - [\theta]_c) ([\theta]_h - [\theta]_c)^{-1} \quad (97)$$

where $[\theta]_c$ and $[\theta]_h$ are the mean residue ellipticities for the completely disordered and completely helical molecule, respectively. A value of $-40,700 \text{ deg cm}^2 \text{ dmol}^{-1}$ was used for $[\theta]_h$. It is the largest negative value measured on our spectropolarimeter for polyhydroxybutyl-glutamine.¹⁸¹ This value is in reasonable agreement with the circular dichroism reported for completely helical poly-L-glutamic acid.²⁰³⁻²⁰⁵ The mean residue ellipticity at 222 nm is vanishingly small for disordered poly-L-glutamic acid, which should permit the approximation of $[\theta]_c$ as zero.

The data from the potentiometric titrations was plotted in the usual manner,^{195,196} and these titration curves are depicted in figures 24 and 25.

Figure 24. The titration curves at 9°, 22°, and 36°C of the L-glutamic acid-L-tyrosine copolymer in 0.1 M NaCl. The value of pK_0 obtained by extrapolation is 4.76 for an L-glutamic acid residue in the copolymer.

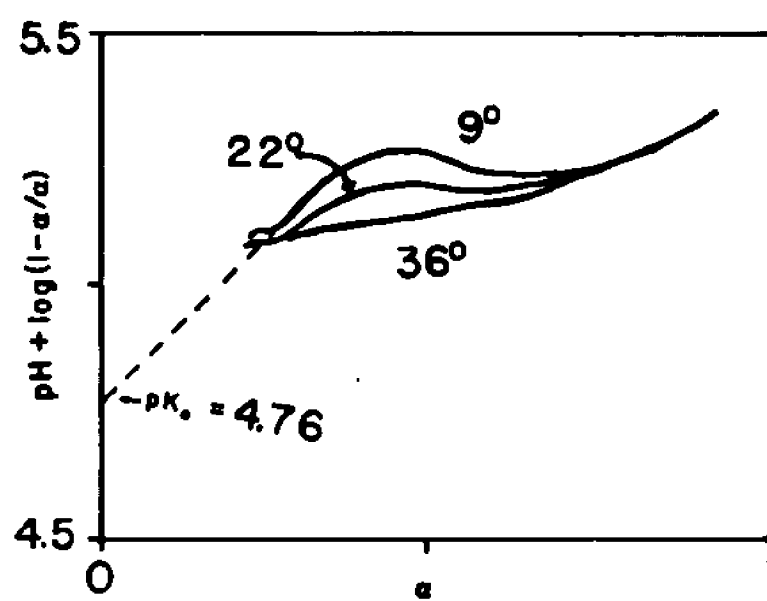
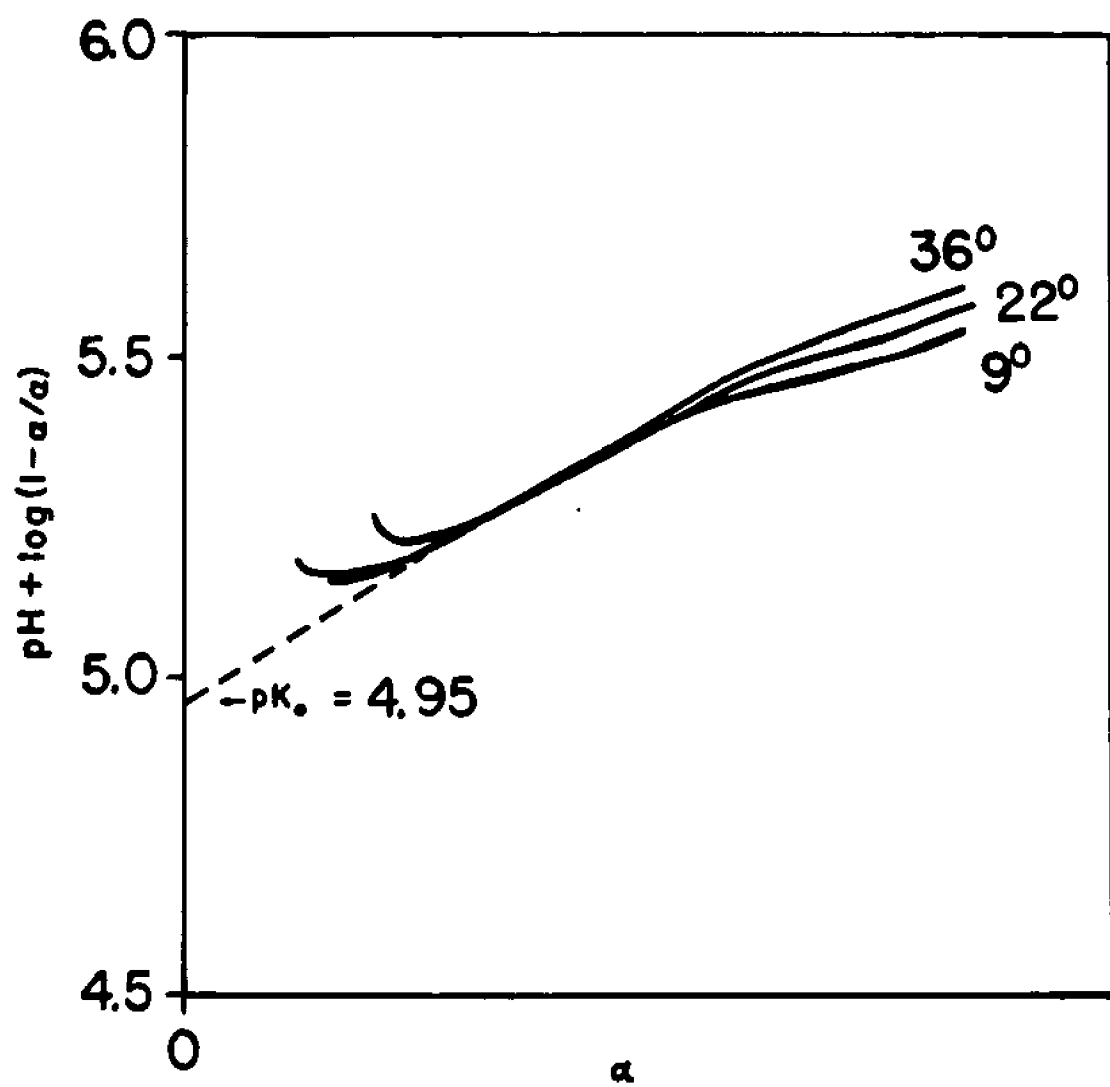


Figure 25. The titration curves at 9°, 22°, and 36°C for the L-glutamic acid-L-alanine copolymer in 0.1 M NaCl. The value for pK_o obtained by extrapolation is 4.95 for an L-glutamic acid residue in the copolymer.



DISCUSSION

The charge induced helix-coil transition in poly-L-glutamic acid has been studied for many years by several groups of researchers!¹⁹⁵⁻²⁰⁰ Several theoretical treatments have been introduced to rationalize this transition. Some of these methods are approximate solutions to rigorous treatments of the electrostatic interactions.¹⁹⁶ The model which shows the most promise for extension to the treatment of the electrostatic interactions in random copolypeptides is that of Zimm and Rice.²⁰⁰ As mentioned previously, this treatment has been applied with good success by Santiago and coworkers¹⁹⁹ to poly-L-glutamic acid at various temperatures and ionic strengths.

One very important consideration for any possible theoretical treatment is whether or not it is actually tractable for the system of interest. If the Zimm-Rice theory is applied, and consideration is given only to nearest neighbor interactions, then the resultant statistical weight matrix is of order 4×4 . If this is combined with methods for obtaining a sampling of computer generated random copolymer molecules, and performing the appropriate calculations, then one has a potentially powerful tool for gaining insight into the nature of the charge induced helix-coil transition in these molecules. This method is certainly not as rigorous as those which are applicable to homopolypeptides, but much can be learned by investigating if this treatment works, and if it does not, why. The parameters used for the L-glutamyl residue are taken from reference 199, and those for the L-alanyl and L-tyrosyl residues are from references 36 and 42, respectively.

The best test of any theoretical treatment is via comparison of the results calculated and those obtained from experiment. A comparison of the results obtained by calculations carried out as described in this chapter and those obtained experimentally is presented in figure 26 for the L-tyrosyl copolymer and figure 27 for the L-alanyl copolymer. The plots depict both the degree of ionization, α , and fraction helix as a function of pH for the copolypeptides in 0.1 M NaCl at 22°C. It can be seen that while the treatment captures the essentials of the transition, the agreement between theory and experiment is not as good as one would hope for. This lack of success is not really surprising considering the fact that only nearest neighbor interactions are being considered. There are several other points to consider before returning to this limitation.

There are several important points to consider with regards to these copolypeptides, which have a bearing on just how good or bad this model is for the transition. One such consideration is whether or not one can justify a different value for the intrinsic dissociation constant for an L-glutamyl residue in the copolypeptides compared with the same residue in the homopolymer. Another point is whether an adjustment of the parameters of the model can be made which would result in better agreement between theory and experiment, and even more important, whether such a change can be justified in terms of the chemistry and physics of the molecules under study. Insight can also be gained by examining what has been done by other workers with similar molecules.

The extrapolation to pK_0 , the intrinsic dissociation constant, is made with the greatest amount of accuracy for these copolymers using the helical portion of the titration curves. It can be seen from the

Figure 26. A plot of the degree of ionization, α , (\blacktriangle) and fraction helix, f_h , (\bullet) as a function of pH for the L-glutamic acid-L-tyrosine copolymer in 0.1 M NaCl at 22°C. The points are from experiment and the lines are calculated from theory.

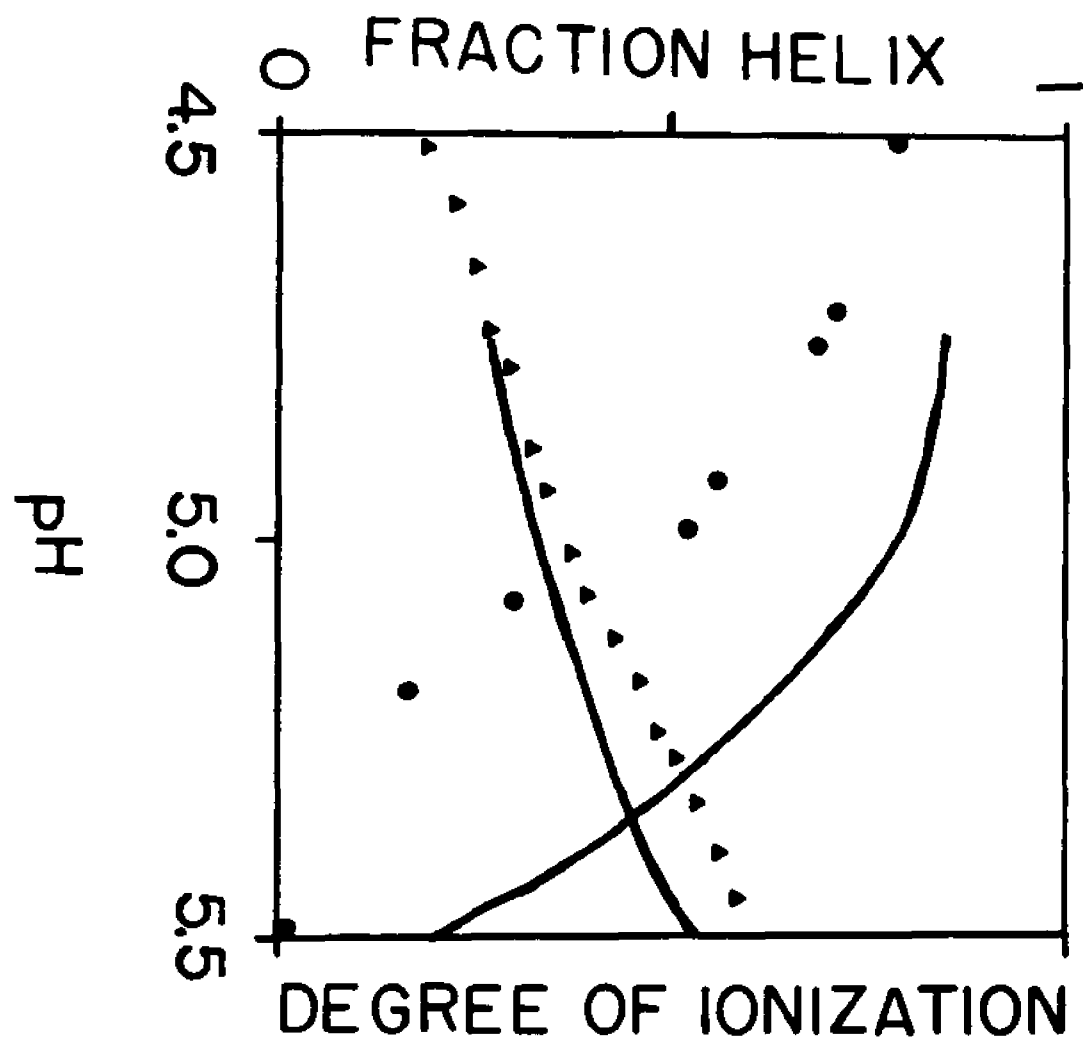
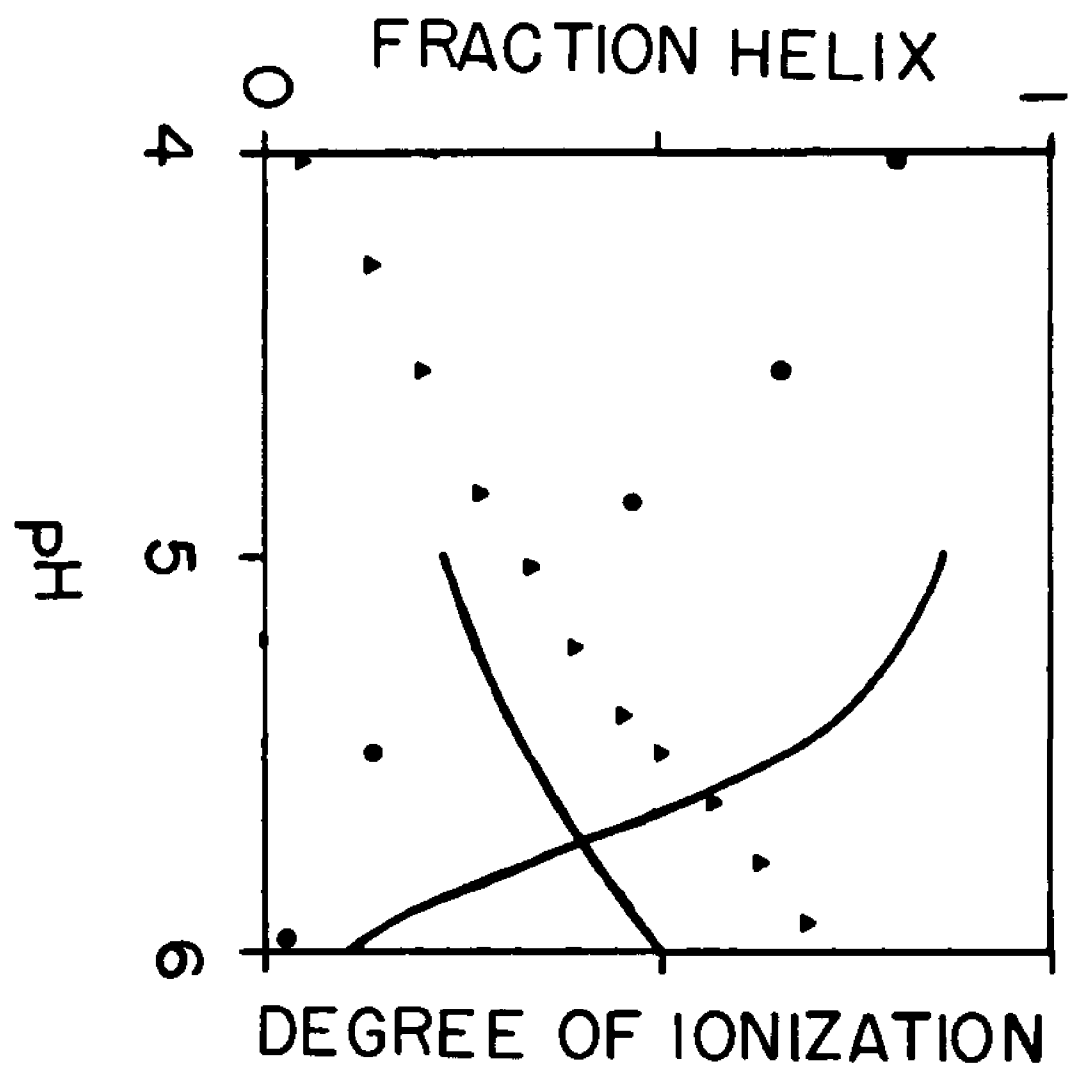


Figure 27. A plot of the degree of ionization, α , (\blacktriangle), and fraction helix, f_h , (\bullet), as a function of pH for the L-glutamic acid-L-alanine copolymer in 0.1 M NaCl at 22°C. The points are from experiment and the lines are calculated from theory.



titration curves presented here, and also those of others who studied random copolypeptides containing L-glutamyl and L-leucyl residues,²⁰⁶ that the extrapolation yields a higher value for pK_0 for the L-glutamyl residues in the copolymer compared to that in the homopolypeptide. This difference appears to increase in proportion to the increase in non-ionizable residue content in the copolypeptides. Nylund and Miller²⁰⁶⁻⁸ dismissed the notion of an increased value for pK_0 for L-glutamyl residues in copolymers containing L-leucyl residues by examination of molecular models. They reasoned that if the access of the sidechain carboxyl groups to the polar solvent is not interfered with or restricted by the nonionizable sidechains, then there is no rationalization for using any value of pK_0 for an L-glutamyl residue other than the one for a residue in the homopolymer. Such conclusions drawn from molecular models are not borne out by the behavior of the molecules in solution, and the experimental results infer that pK_0 is higher.

The next question to address is whether or not the parameters of this model can be adjusted to give better agreement with experiment. The answer to this question is yes, however, the adjustment which has to be made is in conflict with what is known about the physics of these molecules. One must now return to a consideration of the electrostatic interaction parameters which appear in the statistical weight matrices used in the calculations. To be blunt, one needs to consider only one of these parameters, u_{cc} , as the others are considered to be equal to zero.¹⁹⁹ u_{cc} is the contribution to the statistical weight related to the electrostatic interaction between neighboring charged random coil residues. If one examines the differences between the calculated and experimental values for α and f_h as a function of pH,

then it is at once obvious in which direction u_{CC} must be changed in order to maximize agreement between theory and experiment. If the value of u_{CC} is increased then the calculated values of f_h decrease and α increase. This would increase the agreement between theory and experiment, but can such an increase in u_{CC} be justified from fundamental principles ? Strictly speaking, and if only nearest neighbor interactions are considered, then the answer is no. On the average, the charged carboxyls are separated by a greater distance in the copolymers than in the homopolymer, but this effect can not be considered within the context of nearest neighbor interactions. One possible rationalization has to do with the dielectric constant. Clearly, the dielectric constant in the immediate area of the polypeptide backbone is lowered by the incorporation of nonionizable residues into the polymer. By all accounts, however, if anything this would greatly increase the interaction between the two charged groups, and hence result in a lower statistical weight for the two adjacent charged residues in any conformational state. The equations for the electrostatic potential for discrete charges attached to spheres of low dielectric medium immersed in a high dielectric medium have been solved by Tanford and Kirkwood.²⁰⁹ As the charges are embedded further into the low dielectric medium the electrostatic interaction increases enormously. As the charges are further embedded, more of the electric lines of force act through the low dielectric medium and give rise to increased interaction.

This leaves one to ponder the shortcomings of this model which cause the lack of agreement between theory and experiment. Considering the enormous complexity of the system it is easy to think of many reasons for such failure. However, it is of importance here to note

those major contributions to the physics of the system which are ignored in the analysis presented in this chapter, which may possibly be incorporated into a tractable treatment. If one examines the structure of the alpha helix, and especially the disposition of the side-chains, it is at once obvious that interactions between pairs of residues three and four positions apart are as important as those between nearest neighbors. These interactions can be incorporated into a treatment which can ultimately be condensed to a 4×4 statistical weight matrix, and work in this direction should yield valuable information.

In conclusion, it can be seen that the Zimm-Rice theory and the parameters determined from homopolymers does not adequately account for the electrostatic interactions in these random copolypeptides. The reasons have been discussed extensively, and it seems that a path towards a more adequate treatment is both in order and in sight.

BIBLIOGRAPHY

- 1) Anfinsen, C.B. (1973) *Science* 181, 223.
- 2) Linderstrom-Lang, K.U., and Schellman, J.A. *The Enzymes* (P.D. Boyer Editor) Volume 1, second edition, 443. Academic Press, New York, New York 1959.
- 3) Vogel, O., Hoehn, B., and Henning, U. (1972) *Eur. J. Biochem.* 30, 354.
- 4) Pauling, L., Corey, R.B., and Branson, H.R. (1951) *Proc. Natl. Acad. Sci. USA* 37, 205.
- 5) Wieland, T., and Birr, C. Homodetic Cyclic Peptides, *International Review of Science, Organic Series Two*, Volume 6, 183., Butterworth, London 1976.
- 6) Karle, I.L., Karle, J., Wieland, T., Burgermeister, W., Faulstich, H., and Witkop, B. (1973) *Proc. Natl. Acad. Sci. USA* 70, 1836.
- 7) Wycoff, H.W., Tsernoglou, D., Hanson, A.W., Knox, J.R., Lee, B., and Richards, F.M. (1970) *J. Biol. Chem.* 245, 305.
- 8) Schubert-Wright, C., Alden, R.A., and Kraut, J. (1969) *Nature* 221, 235.
- 9) Edsall, J.T., Flory, P.J., Kendrew, J.C., Liquori, A.M., Nemethy, G., Ramachandran, G.N., and Scheraga, H.A. (1966) *J. Mol. Biol.* 15, 399.
- 10) Flory, P.J. *Statistical Mechanics of Chain Molecules* (Wiley-Interscience) New York, New York 1969.
- 11) (1970) *Biochemistry* 9, 3471.
- 12) Ramachandran, G.N., Ramakrishnan, C., and Sasisekharan, U. (1963) *J. Mol. Biol.* 7, 95.
- 13) Brant, D.A., and Flory, P.J. (1965) *J. Am. Chem. Soc.* 87, 2791.
- 14) Brant, D.A., and Flory, P.J. (1965) *J. Am. Chem. Soc.* 87, 663.
- 15) Brant, D.A., Miller, W.G., and Flory, P.J. (1967) *J. Mol. Biol.* 23, 47.
- 16) Ramachandran, G.N., and Sasisekharan, V. (1968) *Adv. Prot. Chem.* 23, 283.
- 17) Jones, J.E. (1924) *Proc. Roy. Soc. London Ser. A* 106, 441.

- 18) Flory, P.J. (1974) *Macromolecules* 7, 381.
- 19) Ooi, T., unpublished results.
- 20) Schimmel, P.R., and Flory, P.J. (1968) *J. Mol. Biol.* 34, 105.
- 21) Mattice, W.L., Nishikawa, K., and Ooi, T. (1973) *Macromolecules* 6, 443.
- 22) Flory, P.J. *Principles of Polymer Chemistry* Cornell U. Press 1953.
- 23) Perutz, M.F. (1951) *Nature* 167, 1053.
- 24) Doty, P., Bradbury, R., Holtzer, A. (1956) *J. Am. Chem. Soc.* 78, 947.
- 25) Doty, P., and Yang, J.T. (1956) *J. Am. Chem. Soc.* 78, 498.
- 26) Schellman, J.R. (1955) *C.R. Trav. Carlsberg Serv. Chim.* 29, 230.
- 27) Poland, D., and Scheraga, H.A. *Theory of Helix-Coil Transitions in Biopolymers*, Academic Press; New York, New York 1970.
- 28) Zimm, B.H., and Bragg, J. (1959) *J. Chem. Phys.* 31, 526.
- 29) Lifson, S., and Roig, A. (1961) *J. Chem. Phys.* 34, 1963.
- 30) Peller, L. (1959) *J. Chem. Phys.* 63, 1149.
- 31) Hill, T.L. (1959) *J. Chem. Phys.* 30, 383.
- 32) Nagai, K. (1960) *J. Phys. Soc. Japan* 15, 407.
- 33) von Dreele, P.H., Lotan, N., Ananthanarayanan, V.S., Andreatta, R.H. Poland, D., and Scheraga, H.A. (1971) *Macromolecules* 4, 408.
- 34) von Dreele, P.H., Poland, D., and Scheraga, H.A. (1971) *Macromolecules* 4, 396.
- 35) Ananthanarayanan, V.S., Andreatta, R.H., Poland, D., and Scheraga, H.A. (1971) *Macromolecules* 4, 417.
- 36) Platzer, K.E.B., Ananthanarayanan, V.S., Andreatta, R.H., and Scheraga, H.A. (1972) *Macromolecules* 5, 177.
- 37) Hughes, L.J., Andreatta, R.H., and Scheraga, H.A. (1972) *Macromolecules* 5, 187.
- 38) Alter, J.E., Taylor, G.T., and Scheraga, H.A. (1972) *Macromolecules* 5, 739.
- 39) van Wart, H.E., Taylor, G.T., and Scheraga, H.A. (1973) *Macromolecules* 6, 266.

- 40) Alter, J.E., Andreatta, R.H., Taylor, G.T., and Scheraga, H.A. (1973) *Macromolecules* 6, 564.
- 41) Maxfield, F.R., Alter, J.E., Taylor, G.T., and Scheraga, H.A. (1975) *Macromolecules* 8, 479.
- 42) Scheule, R.K., Cardinaux, F., Taylor, G.T., and Scheraga, H.A. (1976) *Macromolecules* 9, 23.
- 43) Dygert, M.K., Taylor, G.T., Cardinaux, F., and Scheraga, H.A. (1976) *Macromolecules* 9, 795.
- 44) Kenishi, Y., van Nispen, J.W., Davenport, G., and Scheraga, H.A. (1977) *Macromolecules* 10, 1264.
- 45) Kobayashi, Y., Cardinaux, F., Zweifel, B.O., and Scheraga, H.A. (1977) *Macromolecules* 10, 1271.
- 46) Matheson, R.R., Jr., Nemenhoff, R.A., Cardinaux, F., and Scheraga, H.A. (1977) *Biopolymers* 16, 1567.
- 47) Hill, D.J., Cardinaux, F., and Scheraga, H.A. (1977) *Biopolymers* 16, 2447.
- 48) Hecht, M.H., Zweifel, B.O., and Scheraga, H.A. (1978) *Macromolecules* 11, 545.
- 49) Nagy, J.A., Powers, S.P., Zweifel, B.O., and Scheraga, H.A. (1980) *Macromolecules* 13, 1428.
- 50) Fredrickson, R.A., Chang, M.C., Powers, S.P., and Scheraga, H.A. (1981) *Macromolecules* 14, 625.
- 51) Chang, M.C., Fredrickson, R.A., Powers, S.P., and Scheraga, H.A. (1981) *Macromolecules* 14, 633.
- 52) Jernigan, R.L., as quoted by P.J. Flory and W.G. Miller (1966) *J. Mol. Biol.* 15, 284.
- 53) *Hormonal Peptides and Proteins*, Volume 5, (C.H. Li Editor), 75, Academic Press, New York, New York (1978).
- 54) *Hormonal Peptides and Proteins*, Volume 5, (C.H. Li Editor), 35 Academic Press, New York, New York (1978).
- 55) Wu, C., Lee, N.M., Loh, H.H., Yang, J.T., and Li, C.H. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3656.
- 56) Yang, J.T., Bewley, T.A., Chien, G.C., and Li, C.H. (1977) *Proc. Natl. Acad. Sci. USA* 74, 3235.
- 57) Anselimo, K.J., and Hoffman, E. (1931) *Klin. Wochenschr.* 10, 2280.

- 58) Li, C.H., Barnafi, L., Chretien, M., and Chung, D. (1965) *Nature* 208, 1093.
- 59) Li, C.H., and Chung, D. (1976) *Nature* 260, 622.
- 60) The Mysteries of Opium Revealed, Dr. John Jones, Printed for Richard Smith at the Angel and Bible, London 1701.
- 61) Beckett, A.H. (1959) *Fortschr. Arzneimittelforsch.* 1, 455.
- 62) Goldstein, A., Lowney, L.I., and Pal, B.K. (1971) *Proc. Natl. Acad. Sci. USA* 68, 1742.
- 63) Simon, E.J. (1973) *Proc. Natl. Acad. Sci. USA* 70, 1947.
- 64) Terenius, L. (1973) *Acta Pharmacol. Toxicol.* 33, 377.
- 65) Pert, C.B., and Snyder, S.H. (1973) *Science* 179, 1011.
- 66) Graf, L., Barat, E., Cseh, G., and Sajgo, M. (1971) *Biochim. Biophys. Acta* 229, 276.
- 67) Pert, C.B. (1975) *Brain Research* 75, 112.
- 68) Hughes, J. (1975) *Brain Research* 88, 1.
- 69) Terenius, L., and Wahlstrom, A. (1975) *Life Sci.* 16, 1759.
- 70) Pasternak, G.N., Goodman, R., and Snyder, S.H. (1975) *Life Sci.* 16, 1769.
- 71) Hughes, J., Smith, T.W., Kosterlitz, H.W., Fothergill, L.A., Morgan B.A., and Norris, H.R. (1975) *Nature* 258, 577.
- 72) Li, C.H., Danho, W.O., Chung, D., and Rao, A.J. (1975) *Biochemistry* 14, 497.
- 73) Li, C.H., and Chung, D. (1976) *Proc. Natl. Acad. Sci. USA* 73, 1145.
- 74) Bradbury, A.F., Smyth, D.G., and Snell, C.R. (1976) *Biochem. Biophys. Res. Comm.* 69, 950.
- 75) Bradbury, A.F., Smyth, D.G., and Smell, C.R. (1976) *Nature* 260, 793.
- 76) Graf, L., Barat, E., and Patthy, A. (1976) *Acta Biochim. Biophys. Acad. Sci. Hung.* 11, 121.
- 77) Chretien, M., Benjannet, S., Dragon, N., Seidah, N.G., and Lis, M. (1976) *Biochem. Biophys. Res. Comm.* 72, 472.
- 78) Li, C.H., Lemaire, S., Yamashiro, D., and Doneen, B.A. (1976) *Biochem. Biophys. Res. Comm.* 71, 19.

- 79) Foley, K.M., Kourides, I.A., Inturrisi, C.E., Kaiko, R.F., Zaroulis, C.G., Posner, J.B., Houde, R.W., and Li, C.H. (1979) *Proc. Natl. Acad. Sci. USA* 76, 5377.
- 80) Li, C.H., Tseng, L., and Yamashiro, D. (1978) *Biochem. Biophys. Res. Comm.* 85, 795.
- 81) Blake, J., Tseng, L., Chung, W.C., and Ferrara, P. (1978) *Int. J. Pep. and Prot. Res.* 11, 323.
- 82) Li, C.H., Yamashiro, D., Tseng, L., Chang, W.C., and Ferrara, P. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3276.
- 83) Law, P., Loh, H.H., and Li, C.H. (1979) *Proc. Natl. Acad. Sci. USA* 76, 4255.
- 84) Loh, H.H., Cho, T.M., Wu, Y.C., Harris, R.A., and Way, e.l. (1975) *Life Sci.* 14, 2231.
- 85) Loh, H.H., Cho, T.M., Wu, Y.C., Harris, R.A., and Way, E.L. (1977) *Life Sci.* 16, 1811.
- 86) Loh, H.H., Law, P.Y., Ostwald, T., Cho, T.M., and Way, E.L. (1978) *Fed. Proc. Fed. Am. Soc. Exp. Biol.* 37, 147.
- 87) Hollosi, M., Kajtar, M., and Graf, L. (1977) *F.E.B.S. Lett.* 74, 185.
- 88) Chou, P.Y., and Fasman, G.D. (1974) *Biochemistry* 13, 211.
- 89) Chou, P.Y., and Fasman, G.D. (1974) *Biochemistry* 13, 222.
- 90) Fasman, G.D., Chou, P.Y., and Adler, A.J. (1976) *Biophys. J.* 16, 1201.
- 91) Jibson, M.D., and Li, C.H. (1979) *Int. J. Pep. and Prot. Res.* 14, 113.
- 92) Anthony, J.S., and Moscarello, M.A., (1971) *Biochim. Biophys. Acta* 243, 429.
- 93) Deber, C.M., Moscarello, M.A., and Wood, D.D. (1978) *Biochemistry* 17, 898.
- 94) Keniry, M.A., and Smith, R. (1979) *Biochim. Biophys. Acta* 578, 381.
- 95) Littlemore, L.A.T., and Ledeen, R. (1979) *Aust. J. Chem.* 32, 2631.
- 96) Grouke, M.J., and Gibbs, J.H. (1967) *Biopolymers* 5, 586.
- 97) Satake, I., and Yang, J.T. (1973) *Biochem. Biophys. Res. Comm.* 54, 930.

- 98) Satake, I., and Yang, J.T. (1975) Biopolymers 14, 1841.
- 99) McCord, R.W., Blakeney, E.W., Jr., and Mattice, W.L. (1977) Biopolymers 16, 1329.
- 100) Mattice, W.L., Srinivasan, G., and Santiago, G. (1980) Macromolecules 13, 1254.
- 101) Mattice, W.L., Riser, J.M., and Clark, D.S., (1976) Biochemistry 15, 4264.
- 102) Mattice, W.L. (1979) Biopolymers 18, 225.
- 103) Schimmel, P.R., and Flory, P.J. (1967) Proc. Natl. Acad. Sci. USA 58, 52.
- 104) Li, C.H., Chung, D., and Doneen, B.A. (1976) Biochem. Biophys. Res. Comm. 72, 1542.
- 105) Sarkar, P.K., and Doty, P. (1966) Proc. Natl. Acad. Sci. USA 55, 981.
- 106) Mattice, W.L., and Harrison, W.L. (1976) Biopolymers 15, 559.
- 107) Fasman, G.D., Bodenheimer, E., and Lindblow, C. (1964) Biochemistry 3, 155.
- 108) Lotan, N., Yaron, A., Berger, A., and Sela, M. (1965) Biopolymers 3, 625.
- 109) Igou, D.K., Lo, J.T., Clark, D.S., Mattice, W.L., and Younathan, E.S. (1974) Biochem. Biophys. Res. Comm. 60, 140.
- 110) Krigbaum, W.R., and Hsu, T.S. (1975) Biochemistry 14, 2542.
- 111) Mattice, W.L., and Robinson, R.M. (1981) Biopolymers 20, 1421.
- 112) Li, C.H., Tseng, L., Ferrara, P., and Yamashiro, D. (1980) Proc. Natl. Acad. Sci. USA 77, 2303.
- 113) Cox, B.M., Ross, M., Goldstein, A., and Palmour (1979) Brain Res. 265, 311.
- 114) Ferrara, P., and Li, C.H. (1980) Int. J. Pep. and Prot. Res. 16, 166.
- 115) Bromer, W.W., Sinn, L.G., and Behrens, O.K. (1957) J. Am. Chem. Soc. 79, 2807.
- 116) Mutt, V., Magnusson, S., Jorpes, J.E., and Dahl, E. (1965) Biochemistry 4, 2358.

- 117) Bodansky, M., Ondetti, M.A., Levine, S.D., Narayanan, V.L., von Sultz, M., Sheehan, J.T., Williams, N.J., and Sabo, E.F. (1966) Chem. Ind. 42, 1757.
- 118) Mutt, V., and Said, S.I. (1974) Eur. J. Biochem. 42, 581.
- 119) Bodansky, M., Klausner, Y.S., and Said, S.I. (1973) Proc. Natl. Acad. Sci. USA 70, 382.
- 120) Mutt, V., and Jorpes, J.E. (1967) Rec. Prog. Horm. Res. 23, 483.
- 121) Bodansky, A., Ondetti, M., Mutt, V., and Bodansky, M. (1969) J. Am. Chem. Soc. 91, 944.
- 122) Dockray, G.J. (1979) Fed. Proc. Am. Soc. Exp. Biol. 38, 2295.
- 123) Gardner, J.D., Conlon, T.P., Fink, M.L., and Bodansky, M. (1976) Gastroenterology 71, 965.
- 124) Robberect, P., Conlon, T.P., and Gardner, J.D. (1976) J. Biol. Chem. 251, 4635.
- 125) Mutt, V. (1979) Fed. Proc. Am. Soc. Exp. Biol. 38, 2309.
- 126) Solomon, H.S. (1980) Science 209, 976.
- 127) Robberect, P., Deneef, P., Lammens, M., Deschodt, and Christophe, J.P. (1978) Eur. J. Biochem. 90, 147.
- 128) Taylor, D.P., and Pert, C.B. (1979) Proc. Natl. Acad. Sci. USA 76, 660.
- 129) Robinson, R.M., Hamed, M.M., and Mattice, W.L. (1982) Biochem. Biophys. Res. Comm. 102, 398.
- 130) Mattice, W.L., and Robinson, R.M. (1981) Biochem. Biophys. Res. Comm. 101, 1311.
- 131) Cassim, J.Y., and Yang, J.T. (1970) Biochemistry 8, 1947.
- 132) Gratzer, W.B., Bailey, E., and Beaven, G.H. (1967) Biochem. Biophys. Res. Comm. 28, 914.
- 133) Mattice, W.L., Lo, J.T., and Mandelkern, L. (1972) Macromolecules 5, 729.
- 134) Mattice, W.L., and Harrison, W.H. (1975) Biopolymers 14, 2025.
- 135) Bodansky, A., Ondetti, M.A., and Bodansky, M. (1972) J. Am. Chem. Soc. 94, 3600.
- 136) Bodansky, M., Fink, M.L., Funk, K.W., and Said, S.I. (1976) Clin. Endocrin. (Suppl.) 5, 195.

- 137) Bodansky, M., Bodansky, A., Klausner, Y.S., and Said, S.I. (1974) *Bioorg. Chem.* 3, 133.
- 138) Blanchard, M.H., and King, M.V. (1966) *Biochem. Biophys. Res. Comm.* 25, 298.
- 139) Gratzer, W.B., Beaven, G.H., Rattle, H.W.E., and Bradbury, E.M. (1968) *Eur. J. Biochem.* 3, 276.
- 140) Srerc, P.A., and Brooks, G.C. (1969) *Arch. Biochem. Biophys.* 129, 708.
- 141) Gratzer, W.B., and Beaven, G.H. (1969) *J. Biol. Chem.* 244, 6675.
- 142) Panijpan, B., and Beaven, G.H. (1974) *Eur. J. Biochem.* 45, 547.
- 143) Robinson, R.M., Blakeney, E.W., and Mattice, W.L. (1982) *Biopolymers*, in press.
- 144) Yu, N.T., and Liu, G.S. (1972) *J. Am. Chem. Soc.* 94, 5127.
- 145) Sasaki, K., Dockerill, S., Adamiak, D.A., Tickle, I.J., and Blundell, T. (1975) *Nature* 257, 751.
- 146) Wu, C.C., and Yang, J.T. (1980) *Biochemistry* 19, 2117.
- 147) Robdell, M., Birbbaumer, L., Pohl, S.L., and Sundby, F. (1971) *Proc. Natl. Acad. Sci. USA* 68, 909.
- 148) Chou, P.Y., and Fasman, G.D. (1975) *Biochemistry* 14, 1536.
- 149) Konig, W., Bickel, M., Wissman, H., Uhlmann, R., and Geiger, K. (1979) Hormone Receptors in Digestion and Nutrition, p. 137 Elsevier/North Holland.
- 150) Christophe, J.P., Conlon, T.P., and Gardner, J.D. (1976) *J. Biol. Chem.* 251, 4629.
- 151) De, S.N., and Chatterje, D.N. (1953) *J. Path. Bacteriol.* 6b, 559.
- 152) Carpenter, C.C., Sack, R.B., Feeley, J.C., and Steenburg, R.W. (1968) *J. Clin. Invest.* 47, 1210.
- 153) Field, M., Fromm, D., Wallace, L.K., and Greenough, W.B., III (1969) *J. Clin. Invest.* 48, 24a.
- 154) Vaughn, M., Pierce, N.F., and Greenough, W.B., III (1970) *Nature* 226, 658.
- 155) Chen, L.C., Rohde, J.E., and Sharpe, G.W.G. (1972) *J. Clin. Invest.* 51, 731.
- 156) Sharpe, G.W.G., and Hynie, S. (1971) *Nature* 229, 266.

- 157) Boyle, J.M., and Gardner, J. (1974) J. Clin. Invest. 53, 1149.
- 158) Gill, D.M., and King, C.A. (1975) J. Biol. Chem. 250, 6424.
- 159) Finkelstein, R.A., and LoSpalluto, J.J. (1969) J. Exp. Med. 130, 185.
- 160) Gill, D.M. (1976) Biochemistry 15, 1242.
- 161) Lai, C.Y. (1980) C.R.C. Crit. Rev. in Biochem., 171.
- 162) Wodnar-Filipowicz, A., and Lai, C.Y. (1976) Arch. Biochem. Biophys. 176, 465.
- 163) Cuatrecasas, P. (1973) Biochemistry 12, 3547.
- 164) Cuatrecasas, P. (1973) Biochemistry 12, 3558.
- 165) Cuatrecasas, P. (1973) Biochemistry 12, 3566.
- 166) King, C.A., and van Heyningen, W.E. (1973) J. Infect. Dis. 127, 639.
- 167) Holmgren, J., Lonroth, I., and Svennerholm, L. (1973) Infect. Immun. 8, 208.
- 168) Kurosky, A., Markel, D.E., and Peterson, J.W. (1977) J. Biol. Chem. 252, 7257.
- 169) Duffy, L.K., and Lai, C.Y. (1979) Biochem. Biophys. Res. Comm. 91, 1005.
- 170) Fishman, P.H., Moss, J., and Osborne, J.C. (1978) Biochemistry 17, 711.
- 171) Maroun, R.C., and Mattice, W.L. (1981) Biochem. Biophys. Res. Comm. 103, 442.
- 172) LoSpalluto, J.J., and Finkelstein, R.A. (1972) Biochim. Biophys. Acta 257, 158.
- 173) Zimmerman, S.S., and Scheraga, H.A. (1977) Proc. Natl. Acad. Sci. USA 74, 4126.
- 174) Skolnick, J., and Holtzer, A. (1982) Macromolecules 15, in press.
- 175) Mattice, W.L., and Skolnick, J. (1982) Macromolecules 15, in press.
- 176) Mattice, W.L., and Lo, J.T. (1972) Macromolecules 5, 734.
- 177) Lotan, N., Yaron, A., and Berger, A. (1966) Biopolymers 4, 365.

- 178) Adler, A.J., Hoving, R., Potter, J., Wells, M., and Fasman, G.D. (1968) J. Am. Chem. Soc. 90, 4736.
- 179) Chou, P.Y., Wells, M., and Fasman, G.D. (1971) Biochemistry 11, 3028.
- 180) Okita, K., Teramota, A., and Fujita, H. (1970) Biopolymers 9, 717.
- 181) McCord, R.W., Madison, E.L., and Mattice, W.L. (1980) Biopolymers 19, 1923.
- 182) Yaron, A., Lupu, N., Sela, M., and Berger, A. (1963) Biochim. Biophys. Acta 69, 430.
- 183) Lotan, N., Bixon, M., and Berger, A. (1967) Biopolymers 5, 69.
- 184) Okita, K., Teremoto, A., and Fujitz, H. (1970) Poly. J. 1, 582.
- 185) Inoue, H., and Izumi, T. (1976) Biopolymers 15, 797.
- 186) Feyereisen, C., Morcellet, M., and Loucheux, C. (1977) Macromolecules 11, 620.
- 187) Feyereisen, C., Morcellet, M., and Loucheux, C. (1978) Macromolecules 11, 620.
- 188) McCord, R.W. (1979) Ph.D. Dissertation, Louisiana State University, Baton Rouge.
- 189) Pym, R.V.E., and Milham, P.J. (1976) Anal. Chem. 48, 1413.
- 190) Lifson, S. (1963) Biopolymers 1, 25.
- 191) Allegra, G. (1967) J. Polym. Sci., Part C 16, 2815.
- 192) Denton, J.B., Powers, S.P., Zweifel, B.O., and Scheraga, H.A. (1982) Biopolymers 21, 51.
- 193) Miyake, M., Akita, S., Teramoto, A., Norisyue, T., and Fumita, H. Biopolymers 13, 1173.
- 194) Paterson, Y., and Leach, S.J. (1978) Macromolecules 11, 418.
- 195) Wada, A. (1960) Mol. Phys. 3, 409.
- 196) Nagasawa, M., and Holtzer, A. (1964) J. Am. Chem. Soc. 86, 538.
- 197) Olander, D.S., and Holtzer, A. (1968) J. Am. Chem. Soc. 90, 4549.
- 198) Warashina, A., and Ikegama, A. (1972) Biopolymers 11, 529.
- 199) Santiago, G., Maroun, R.C., Hawkins, E.R., and Mattice, W.L. (1981) Biopolymers 20, 2181.

- 200) Zimm, B.H., and Rice, S.A. (1960) Mol. Phys. 3, 391.
- 201) Rice, S.A., and Nagasawa, M. Polyelectrolyte Solutions : A Theoretical Introduction, Academic Press ; New York, New York(1961).
- 202) D. Olander, Ph.D. Thesis, Washington University, St. Louis, Mo, (1967).
- 203) Holzwarth, G., and Doty, P. (1965) J. Am. Chem. Soc. 87, 218.
- 204) Hammes, G.G., and Schullery, S.E. (1968) Biochemistry 7, 3882.
- 205) Cassim, J.Y., and Yang, J.T. (1970) Biopolymers 9, 1475.
- 206) Nylund, R.E., and Miller, W.G. (1965) J. Am. Chem. Soc. 87, 3537.
- 207) Miller, W.G., and Nylund, R.E. (1965) J. Am. Chem. Soc. 87, 3547.
- 208) Snipp, R.L., Miller, W.G., and Nylund, R.E. (1965) J. Am. Chem. Soc. 87, 3557.
- 209) Tanford, C., and Kirkwood, J. (1957) J. Am. Chem. Soc. 79, 5333.

VITA

Randall Murdock Robinson was born in Nashville, Tennessee in 1955. He was graduated from Glenclyff High School in Nashville in 1973. He graduated with a Bachelor of Science in Chemistry from Middle Tennessee State University in 1977. He is presently a candidate for the degree of Doctor of Philosophy with a major in Biochemistry and a minor in Chemistry.

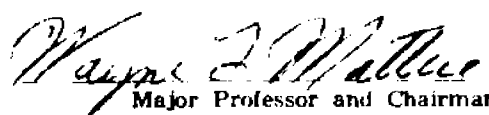
EXAMINATION AND THESIS REPORT


Candidate: Randall M. Robinson

Major Field: Biochemistry

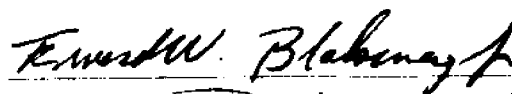
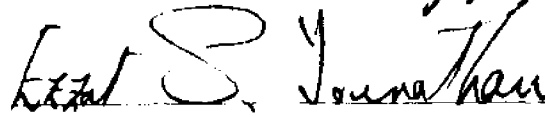
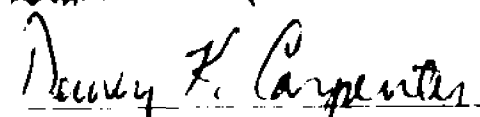

Title of Thesis: Studies on Several Biologically Active Peptides and Synthetic
Random Copolypeptides

Approved:


Major Professor and Chairman


Dean of the Graduate School

EXAMINING COMMITTEE:

Date of Examination: April 14, 1982